

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日
Date of Application:

2002年 9月19日

出 願 番 号
Application Number:

特願2002-273162

[ST.10/C]:

[JP 2002-273162]

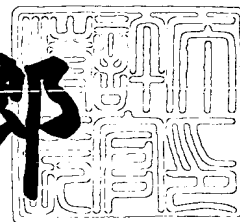
出 願 人
Applicant(s):

株式会社日立製作所

2003年 6月 6日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3044280

【書類名】 特許願

【整理番号】 HI020540

【提出日】 平成14年 9月19日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/00

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 R A I D システム事業部内

【氏名】 金井 宏樹

【発明者】

【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社日立製作所 R A I D システム事業部内

【氏名】 金子 誠司

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100071283

【弁理士】

【氏名又は名称】 一色 健輔

【選任した代理人】

【識別番号】 100084906

【弁理士】

【氏名又は名称】 原島 典孝

【選任した代理人】

【識別番号】 100098523

【弁理士】

【氏名又は名称】 黒川 恵

【選任した代理人】

【識別番号】 100112748

【弁理士】

【氏名又は名称】 吉田 浩二

【選任した代理人】

【識別番号】 100110009

【弁理士】

【氏名又は名称】 青木 康

【手数料の表示】

【予納台帳番号】 011785

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶制御装置、記憶システム、記憶制御装置の制御方法、チャンネル制御部、及びプログラム

【特許請求の範囲】

【請求項 1】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、

前記キャッシュメモリが、前記キャッシュメモリに記憶されたデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されていることを特徴とする記憶制御装置。

【請求項 2】 情報処理装置とのインタフェースを有するチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、

第 1 のキャッシュメモリが前記チャンネル制御部に配置され、第 2 のキャッシュメモリが前記内部接続部に接続されていることを特徴とする記憶制御装置。

【請求項 3】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、

第 1 のキャッシュメモリが、前記第 1 のキャッシュメモリに記憶されたデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置され、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリが、前記内部接続部に接続されていることを特徴とする記憶制御装置。

【請求項 4】 請求項 1 乃至請求項 3 に記載の記憶制御装置と、前記記憶装

置とを備えることを特徴とする記憶システム。

【請求項 5】 情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されているキャッシュメモリと、前記チャネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置に前記情報処理装置からデータの書き込み要求があった場合における、前記チャネル制御部による前記キャッシュメモリへの前記データの書き込み方法であって、

前記情報処理装置から書き込みデータを受信するステップと、

前記書き込みデータを前記キャッシュメモリに書き込むステップと、

前記専用のデータ転送路で相互に結合された他方のチャネル制御部に対して、前記専用のデータ転送路を介して前記書き込みデータを送信するステップと、

前記他方のチャネル制御部から、前記他方のチャネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を前記専用のデータ転送路を介して受信するステップと、

前記情報処理装置に対して、前記書き込みデータの前記キャッシュメモリへの書き込みが完了した旨の報告を送信するステップと
を備えることを特徴とするキャッシュメモリへのデータの書き込み方法。

【請求項 6】 情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部による、前記第 2 のキャッシュメモリに記憶されているデータの第 1 のキャッシュメモリへ

の読み込み方法であって、

前記第 2 のキャッシュメモリへ前記データの読み込みコマンドを送信するステップと、

前記第 2 のキャッシュメモリから前記データを取得するステップと、

前記取得したデータを前記第 1 のキャッシュメモリに書き込むステップと、

前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信するステップと、

前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップと

を備えることを特徴とするキャッシュメモリへのデータの読み込み方法。

【請求項 7】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置に前記情報処理装置からデータの読み出し要求があった場合における、前記チャンネル制御部による前記データの読み出し方法であって、

前記情報処理装置から前記データのアドレスが指定された読み出しコマンドを受信するステップと、

前記指定されたアドレスのデータが前記第 1 のキャッシュメモリに記憶されているかどうかを判定するステップと、

前記指定されたアドレスのデータが前記第 1 のキャッシュメモリに記憶されていない場合には、前記第 2 のキャッシュメモリへ前記データの読み込みコマンドを送信するステップと、

前記第 2 のキャッシュメモリから前記データを取得するステップと、

前記取得したデータを前記第 1 のキャッシュメモリに書き込むステップと、
前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、
前記専用のデータ転送路により前記取得したデータを送信するステップと、

前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップと、

前記情報処理装置に対して前記データを送信するステップと
を備えることを特徴とするデータの読み出し方法。

【請求項 8】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部による、前記第 1 のキャッシュメモリに記憶されているデータの前記第 2 のキャッシュメモリへの書き出し方法であって、

前記第 1 のキャッシュメモリに記憶されているデータの中から前記第 2 のキャッシュメモリへ書き出すデータを選定するステップと、

前記第 2 のキャッシュメモリに前記選定したデータを書き込むための記憶領域を確保するステップと、

前記選定したデータを前記第 2 のキャッシュメモリへ送信するステップと
を備えることを特徴とする第 2 のキャッシュメモリへのデータの書き出し方法。

【請求項 9】 前記選定は最も長期間アクセスのなかったデータを選出することにより行われることを特徴とする請求項 8 に記載のデータの書き出し方法。

【請求項 10】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記

憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置への電力供給が停止した場合における、前記記憶制御装置による前記第1のキャッシュメモリ及び前記第2のキャッシュメモリに記憶されているデータの前記記憶装置への保存方法であって、

前記記憶制御装置にあらかじめ備えられたバッテリーに電力供給源を切り替えるステップと、

前記第1のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記第2のキャッシュメモリに未書き込みのものを前記第2のキャッシュメモリへ書き込むステップと、

前記第2のキャッシュメモリへの書き込み完了後に、前記バッテリーから前記チャンネル制御部への供給電力を低減するステップと、

前記第2のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記記憶装置に未書き込みのものを前記記憶装置へ書き込むステップと、

前記記憶装置への書き込み完了後に、前記バッテリーから前記記憶制御装置への電力供給を遮断するステップと

を備えることを特徴とする記憶装置へのデータの保存方法。

【請求項11】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されているキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部であって、

前記情報処理装置から書き込みデータを受信する手段と、

前記書き込みデータを前記キャッシュメモリに書き込む手段と、

前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、
前記専用のデータ転送路を介して前記書き込みデータを送信する手段と、

前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を前記専用のデータ転送路を介して受信する手段と、

前記情報処理装置に対して、前記書き込みデータの前記キャッシュメモリへの書き込みが完了した旨の報告を送信する手段と
を備えることを特徴とするチャンネル制御部。

【請求項 1 2】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部であって、

前記第 2 のキャッシュメモリに記憶されているデータの読み込みコマンドを前記第 2 のキャッシュメモリに送信する手段と、

前記データを前記第 2 のキャッシュメモリから取得する手段と、

前記取得したデータを前記第 1 のキャッシュメモリに書き込む手段と、

前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、
前記専用のデータ転送路により前記取得したデータを送信する手段と、

前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信する手段と

を備えることを特徴とするチャンネル制御部。

【請求項 1 3】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記

憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部であって、

前記情報処理装置からアドレスが指定されたデータの読み出しコマンドを受信する手段と、

前記指定されたアドレスのデータが前記第1のキャッシュメモリに記憶されているかどうかを判定する手段と、

前記指定されたアドレスのデータが前記第1のキャッシュメモリに記憶されていない場合には、前記第2のキャッシュメモリへ前記データの読み込みコマンドを送信する手段と、

前記第2のキャッシュメモリから前記データを取得する手段と、

前記取得したデータを前記第1のキャッシュメモリに書き込む手段と、

前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信する手段と、

前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信する手段と、

前記情報処理装置に対して前記データを送信する手段とを備えることを特徴とするチャンネル制御部。

【請求項14】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部であって、

前記第 1 のキャッシュメモリに記憶されているデータの中から前記第 2 のキャッシュメモリへ書き出すデータを選定する手段と、

前記第 2 のキャッシュメモリに前記選定したデータを書き込むための記憶領域を確保する手段と、

前記選定したデータを前記第 2 のキャッシュメモリへ送信する手段とを備えることを特徴とするチャンネル制御部。

【請求項 1 5】 前記選定は最も長期間アクセスのなかったデータを選出することにより行われることを特徴とする請求項 1 4 に記載のチャンネル制御部。

【請求項 1 6】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置であって、

前記記憶制御装置への電力供給が停止した場合に、継続して電力供給を行うためのバッテリーと、

前記記憶制御装置への電力供給が停止した場合に、前記バッテリーに電力供給源を切り替える手段と、

前記第 1 のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記第 2 のキャッシュメモリに未書き込みのものを前記第 2 のキャッシュメモリへ書き込む手段と、

前記第 2 のキャッシュメモリへの書き込み完了後に、前記バッテリーから前記チャンネル制御部への供給電力を低減する手段と、

前記第 2 のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記記憶装置に未書き込みのものを前記記憶装置へ書き込む手段と、

前記記憶装置への書き込み完了後に、前記バッテリーから前記記憶制御装置への

電力供給を遮断する手段と

を備えることを特徴とする記憶制御装置。

【請求項 1 7】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されているキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部に、

前記情報処理装置から書き込みデータを受信するステップと、

前記書き込みデータを前記キャッシュメモリに書き込むステップと、

前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路を介して前記書き込みデータを送信するステップと、

前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を前記専用のデータ転送路を介して受信するステップと、

前記情報処理装置に対して、前記書き込みデータの前記キャッシュメモリへの書き込みが完了した旨の報告を送信するステップと
を実行させるためのプログラム。

【請求項 1 8】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部に、

前記第 2 のキャッシュメモリに記憶されているデータの読み込みコマンドを前記第 2 のキャッシュメモリに送信するステップと、

前記データを前記第 2 のキャッシュメモリから取得するステップと、
 前記取得したデータを前記第 1 のキャッシュメモリに書き込むステップと、
 前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、
 前記専用のデータ転送路により前記取得したデータを送信するステップと、
 前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップと
 を実行させるためのプログラム。

【請求項 19】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部に、

前記情報処理装置からアドレスが指定されたデータの読み出しコマンドを受信するステップと、

前記指定されたアドレスのデータが前記第 1 のキャッシュメモリに記憶されているかどうかを判定するステップと、

前記指定されたアドレスのデータが前記第 1 のキャッシュメモリに記憶されていない場合には、前記第 2 のキャッシュメモリへ前記データの読み込みコマンドを送信するステップと、

前記第 2 のキャッシュメモリから前記データを取得するステップと、
 前記取得したデータを前記第 1 のキャッシュメモリに書き込むステップと、
 前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、
 前記専用のデータ転送路により前記取得したデータを送信するステップと、
 前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステ

ップと、

前記情報処理装置に対して前記データを送信するステップと
を実行させるためのプログラム。

【請求項 2 0】 情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部に、

前記第 1 のキャッシュメモリに記憶されているデータの中から前記第 2 のキャッシュメモリへ書き出すデータを選定するステップと、

前記第 2 のキャッシュメモリに前記選定したデータを書き込むための記憶領域を確保するステップと、

前記選定したデータを前記第 2 のキャッシュメモリへ送信するステップと
を実行させるためのプログラム。

【請求項 2 1】 前記選定は最も長期間アクセスのなかったデータを選出することにより行われることを特徴とする請求項 2 0 に記載のプログラム。

【請求項 2 2】 情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置に、

前記記憶制御装置への電力供給が停止した場合に継続して電力供給を行うためのバッテリーに電力供給源を切り替えるステップと、

前記第 1 のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記第 2 のキャッシュメモリに未書き込みのものを前記第 2 のキャッシュメモリへ書き込むステップと、

前記第 2 のキャッシュメモリへの書き込み完了後に、前記バッテリーから前記チャンネル制御部への供給電力を低減するステップと、

前記第 2 のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記記憶装置に未書き込みのものを前記記憶装置へ書き込むステップと、

前記記憶装置への書き込み完了後に、前記バッテリーから前記記憶制御装置への電力供給を遮断するステップと
を実行させるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、記憶制御装置、記憶システム、記憶制御装置の制御方法、チャンネル制御部、及びプログラムに関する。

【0002】

【従来の技術】

従来から、磁気ディスクを記録媒体とする大容量記憶装置のアクセス速度を向上させるために様々な工夫がなされている。その一つとして記憶装置へのアクセスを制御する記憶制御装置の内部に半導体を記録媒体とするキャッシュメモリを置くことが行われている。これにより磁気ディスクから一度データを読み出した後はキャッシュメモリに対して読み書きを行えるので、アクセス速度を向上させることができる。

従来の記憶制御装置では、ホスト計算機との入出力制御を行うチャンネル制御部と、記憶装置との入出力制御を行うディスク制御部との間をネットワークで接続し、ネットワーク上にキャッシュメモリを配置する構成がある。また、キャッシュメモリはデータの消失に備えて 2 重化されていることが多い。

【0003】

しかしこのような構成では、計算機システムの規模が増大し、接続されるホスト計算機や記憶装置内部の磁気ディスクの数が増加した場合に、キャッシュメモリへのデータアクセスが集中し性能向上を妨げる要因となる。

これを回避するためにはキャッシュメモリへのアクセスパスを高帯域化することが有効であるが、高性能なハードウェアの採用が必要となりコストがアップする。またシステムの規模が大きくなる程、より高帯域なアクセスパスが必要となる。

このような問題を回避する技術として、キャッシュメモリを分散して配置することが提案されている（例えば、特許文献 1 参照）。

【 0 0 0 4 】

【特許文献 1】

特開平 1 1 - 2 0 3 2 0 1 号公報

【 0 0 0 5 】

【発明が解決しようとする課題】

しかしながら前述した従来の技術にあっては、キャッシュメモリの分散配置によりキャッシュメモリへのアクセス集中は低減できるものの、データパススイッチ（チャンネル制御部とディスク制御部とを接続するネットワーク）を経由してキャッシュメモリへのアクセスが行われるため、多数のアクセスが同時に発生した場合、データパススイッチ上の経路が競合し、アクセス性能の向上を妨げる要因となる。

また、チャンネル制御部とディスク制御部とを接続するネットワーク上に複数のキャッシュメモリを分散配置する構成では、キャッシュメモリ間でデータを 2 重化して記憶するために行われるデータ転送はネットワークを介して行われるため、ネットワークの使用率が上昇し、ネットワーク上のデータ競合が発生しやすくなる。

そこで、本発明はデータのアクセス高速化を可能とする記憶制御装置を提供することを主たる目的とする。

【 0 0 0 6 】

【課題を解決するための手段】

情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置における、前記キャッシュメモリの配置方法であって、前記キャッシュメモリに記憶されたデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に前記キャッシュメモリを配置する。

【 0 0 0 7 】

その他、本願が開示する課題、及びその解決方法は、発明の実施の形態の欄、実施例の欄、及び図面により明らかにされる。

【 0 0 0 8 】

【発明の実施の形態】

本明細書の記載により少なくとも次のことが明らかにされる。

本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、前記キャッシュメモリが、前記キャッシュメモリに記憶されたデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されていることを特徴とする。

【 0 0 0 9 】

このような態様により、キャッシュメモリ間でのデータ2重化のためのデータ転送は内部接続部を使用せずに行われるので、内部接続部の使用率を抑制することができる。そのため内部接続部でのデータ競合が減少し、情報処理装置へのデータアクセスを高速化できる。またデータの2重化のためにペアを組んでいるキャッシュメモリ間は専用のデータ転送路により結合されているので、データの2重化処理を短時間で行うことができる。これにより情報処理装置に対してデータ

アクセスの完了報告を短時間で行うことができ、データアクセスの高速化を図ることができる。またペアを組むチャンネル制御部の間で相互にキャッシュメモリのデータを2重化することにより、記憶制御装置の運転に影響を与えずに、チャンネル制御部の保守・管理を行うことができる。例えば、あるチャンネル制御部のキャッシュメモリに記憶されているデータは、ペアを組むチャンネル制御部にも記憶されているため、当該チャンネル制御部のキャッシュメモリのデータを当該チャンネル制御部以外の他の記憶領域に待避させるようなことを行わなくとも、当該チャンネル制御部の交換を行うことができる。これにより、24時間365日の連続運転が要求される記憶制御装置の運転を停止させたり機能を低下させたりすることなく、保守・管理作業を行うことが可能となる。またキャッシュメモリがチャンネル制御部に配置されることにより、情報処理装置からアクセス要求のあったデータがキャッシュメモリに存在すれば、情報処理装置は当該データに高速にアクセス可能となる。またこのアクセスは内部接続部を経由しないので、内部接続部でのデータ競合を減少できる。

【0010】

また本発明の一態様は、情報処理装置とのインタフェースを有するチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、第1のキャッシュメモリが前記チャンネル制御部に配置され、第2のキャッシュメモリが前記内部接続部に接続されていることを特徴とする。

【0011】

このような態様により、情報処理装置からのアクセスデータが第1のキャッシュメモリに記憶されていなかった場合でも、第2のキャッシュメモリに当該データが記憶されていれば記憶装置にまでアクセスをする必要が無い。そのため、情報処理装置からのデータのアクセス速度を高速化することが可能となる。

【0012】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャ

ネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、第1のキャッシュメモリが、前記第1のキャッシュメモリに記憶されたデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置され、相互に同一のデータを記憶する2つの第2のキャッシュメモリが、前記内部接続部に接続されていることを特徴とする。

【 0 0 1 3 】

このような態様により、第1のキャッシュメモリ間でのデータ2重化のためのデータ転送は内部接続部を使用せずに行われるので、内部接続部の使用率を抑制することができる。そのため内部接続部でのデータ競合が減少し、情報処理装置へのデータアクセスを高速化できる。またデータの2重化のためにペアを組んでいる第1のキャッシュメモリ間は専用のデータ転送路により結合されているので、データの2重化処理を短時間で行うことができる。これにより情報処理装置に対してデータアクセスの完了報告を短時間で行うことができ、データアクセスの高速化を図ることができる。またペアを組むチャンネル制御部の間で相互にキャッシュメモリのデータを2重化することにより、記憶制御装置の運転に影響を与えずに、チャンネル制御部の保守・管理を行うことができる。例えば、あるチャンネル制御部のキャッシュメモリに記憶されているデータは、ペアを組むチャンネル制御部にも記憶されているため、当該チャンネル制御部のキャッシュメモリのデータを当該チャンネル制御部以外の他の記憶領域に待避させるようなことを行わなくとも、当該チャンネル制御部の交換を行うことができる。これにより、24時間365日の連続運転が要求される記憶制御装置の運転を停止させたり機能を低下させたりすることなく、保守・管理作業を行うことが可能となる。また第1のキャッシュメモリがチャンネル制御部に配置されることにより、情報処理装置からアクセス要求のあったデータが第1のキャッシュメモリに存在すれば、情報処理装置は当該データに高速にアクセス可能となる。またこのアクセスは内部接続部を経由しないので、内部接続部でのデータ競合を減少できる。また、情報処理装置からの

アクセスデータが第1のキャッシュメモリに記憶されていなかった場合でも、第2のキャッシュメモリに当該データが記憶されていれば記憶装置にまでアクセスをする必要が無い。そのため、情報処理装置からのデータのアクセス速度を高速化することが可能となる。

【0014】

また本発明の一態様は、前記記憶制御装置と、前記記憶装置とを備えることを特徴とする。

【0015】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されているキャッシュメモリと、前記チャネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置に前記情報処理装置からデータの書き込み要求があった場合における、前記チャネル制御部による前記キャッシュメモリへの前記データの書き込み方法であって、前記情報処理装置から書き込みデータを受信するステップと、前記書き込みデータを前記キャッシュメモリに書き込むステップと、前記専用のデータ転送路で相互に結合された他方のチャネル制御部に対して、前記専用のデータ転送路を介して前記書き込みデータを送信するステップと、前記他方のチャネル制御部から、前記他方のチャネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を前記専用のデータ転送路を介して受信するステップと、前記情報処理装置に対して、前記書き込みデータの前記キャッシュメモリへの書き込みが完了した旨の報告を送信するステップとを備えることを特徴とする。

【0016】

このような態様により、キャッシュメモリ間でのデータ2重化のためのデータ転送は内部接続部を使用せずに行われるので、内部接続部の使用率を抑制することができる。そのため内部接続部でのデータ競合が減少し、情報処理装置へのデ

ータアクセスを高速化できる。またデータの2重化のためにペアを組んでいるキャッシュメモリ間は専用のデータ転送路により結合されているので、データの2重化処理を短時間で行うことができる。これにより情報処理装置に対してデータアクセスの完了報告を短時間で行うことができ、データアクセスの高速化を図ることができる。またペアを組むチャンネル制御部の間で相互にキャッシュメモリのデータを2重化することにより、記憶制御装置の運転に影響を与えずに、チャンネル制御部の保守・管理を行うことができる。例えば、あるチャンネル制御部のキャッシュメモリに記憶されているデータは、ペアを組むチャンネル制御部にも記憶されているため、当該チャンネル制御部のキャッシュメモリのデータを当該チャンネル制御部以外の他の記憶領域に待避させるようなことを行わなくとも、当該チャンネル制御部の交換を行うことができる。これにより、24時間365日の連続運転が要求される記憶制御装置の運転を停止させたり機能を低下させたりすることなく、保守・管理作業を行うことが可能となる。

【 0 0 1 7 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部による、前記第2のキャッシュメモリに記憶されているデータの前記第1のキャッシュメモリへの読み込み方法であって、前記第2のキャッシュメモリへ前記データの読み込みコマンドを送信するステップと、前記第2のキャッシュメモリから前記データを取得するステップと、前記取得したデータを前記第1のキャッシュメモリに書き込むステップと、前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信するステップと、前記他方のチャンネル制御部から、前記他方のチャンネル制御部が

備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップとを備えることを特徴とする。

【 0 0 1 8 】

このような態様により、第1のキャッシュメモリ間のデータの2重化は専用のデータ転送路を介することにより行われ、内部接続部を使用しないので、内部接続部の使用率を抑制することができる。そのため内部接続部でのデータ競合が減少し、情報処理装置へのデータアクセスを高速化できる。また第2のキャッシュメモリから読み出したデータは2つの第1のキャッシュメモリに2重化されて書き込まれるのでデータの信頼性を向上させることができる。またペアを組むチャンネル制御部の間で相互にキャッシュメモリのデータを2重化することにより、記憶制御装置の運転に影響を与えずに、チャンネル制御部の保守・管理を行うことができる。例えば、あるチャンネル制御部のキャッシュメモリに記憶されているデータは、ペアを組むチャンネル制御部にも記憶されているため、当該チャンネル制御部のキャッシュメモリのデータを当該チャンネル制御部以外の他の記憶領域に待避させるようなことを行わなくとも、当該チャンネル制御部の交換を行うことができる。これにより、24時間365日の連続運転が要求される記憶制御装置の運転を停止させたり機能を低下させたりすることなく、保守・管理作業を行うことが可能となる。

【 0 0 1 9 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置に前記情報処理装置からデータの読み出し要求があった場合における、前記チャンネル制御部による前記データの読み出し方法であって、前記情報処理装置から前記データのアドレスが指定された読み

出しコマンドを受信するステップと、前記指定されたアドレスのデータが前記第 1 のキャッシュメモリに記憶されているかどうかを判定するステップと、前記指定されたアドレスのデータが前記第 1 のキャッシュメモリに記憶されていない場合には、前記第 2 のキャッシュメモリへ前記データの読み込みコマンドを送信するステップと、前記第 2 のキャッシュメモリから前記データを取得するステップと、前記取得したデータを前記第 1 のキャッシュメモリに書き込むステップと、前記専用のデータ転送路で相互に結合された他方のチャネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信するステップと、前記他方のチャネル制御部から、前記他方のチャネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップと、前記情報処理装置に対して前記データを送信するステップとを備えることを特徴とする。

【 0 0 2 0 】

このような態様により、情報処理装置からデータのリード要求を受けた場合、当該データが第 1 のキャッシュメモリに記憶されていなかった場合でも、第 2 のキャッシュメモリに記憶されていれば、記憶装置にまでアクセスをすることなく当該データを情報処理装置に送信できる。また第 2 のキャッシュメモリから読み出したデータは 2 つの第 1 のキャッシュメモリに 2 重化されて書き込まれる。そのため次回以降のアクセスに対しては第 1 のキャッシュメモリ上のデータで対応することが可能となる。また第 1 のキャッシュメモリ間のデータの 2 重化は専用のデータ転送路を介することにより行われ、内部接続部を使用しないので、内部接続部の使用率を抑制することができる。そのため内部接続部でのデータ競合が減少し、情報処理装置へのデータアクセスを高速化できる。また第 1 のキャッシュメモリ上のデータは 2 重化されて記憶されているので、データの信頼性を向上させることができる。またこの 2 重化はペアを組んでいるチャネル制御部の間で行われているため、記憶制御装置の運転に影響を与えずに、チャネル制御部の保守・管理を行うことができる。例えば、あるチャネル制御部のキャッシュメモリに記憶されているデータは、ペアを組むチャネル制御部にも記憶されているため、当該チャネル制御部のキャッシュメモリのデータを当該チャネル制御部以外の

他の記憶領域に待避させるようなことを行わなくとも、当該チャネル制御部の交換を行うことができる。これにより、24時間365日の連続運転が要求される記憶制御装置の運転を停止させたり機能を低下させたりすることなく、保守・管理作業を行うことが可能となる。

【0021】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部による、前記第1のキャッシュメモリに記憶されているデータの第2のキャッシュメモリへの書き出し方法であって、前記第1のキャッシュメモリに記憶されているデータの中から前記第2のキャッシュメモリへ書き出すデータを選定するステップと、前記第2のキャッシュメモリに前記選定したデータを書き込むための記憶領域を確保するステップと、前記選定したデータを前記第2のキャッシュメモリへ送信するステップとを備えることを特徴とする。

【0022】

このような態様により、第1のキャッシュメモリ上の記憶領域を解放することが必要になった場合、第1のキャッシュメモリからリプレイスされるデータの第2のキャッシュメモリへの書き込みが完了すれば、当該データが記憶されていた第1のキャッシュメモリ上の記憶領域を解放することができるようになる。従来は第2のキャッシュメモリが無かったため、リプレイスされるデータが記憶装置に書き込まれるまで第1のキャッシュメモリの記憶領域を解放することができなかった。本実施の態様により短時間でキャッシュメモリのリプレイス処理を行うことが可能となる。

【0023】

また本発明の一態様は、前記選定は最も長期間アクセスのなかったデータを選出することにより行われることを特徴とする。

このような態様により、第1のキャッシュメモリへ記憶しておく必要性の最も少ないと推定されるデータを第2のキャッシュメモリへリプレースすることが可能となる。

【0024】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置への電力供給が停止した場合における、前記記憶制御装置による前記第1のキャッシュメモリ及び前記第2のキャッシュメモリに記憶されているデータの前記記憶装置への保存方法であって、前記記憶制御装置にあらかじめ備えられたバッテリーに電力供給源を切り替えるステップと、前記第1のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記第2のキャッシュメモリに未書き込みのものを前記第2のキャッシュメモリへ書き込むステップと、前記第2のキャッシュメモリへの書き込み完了後に、前記バッテリーから前記チャネル制御部への供給電力を低減するステップと、前記第2のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記記憶装置に未書き込みのものを前記記憶装置へ書き込むステップと、前記記憶装置への書き込み完了後に、前記バッテリーから前記記憶制御装置への電力供給を遮断するステップとを備えることを特徴とする。

このような態様により、突然の停電等が発生してもキャッシュメモリ上のデータを消失させることがなく、システムの信頼性を向上させることが可能となる。

【0025】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されているキャッシュメモリと、前記チャネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部であって、前記情報処理装置から書き込みデータを受信する手段と、前記書き込みデータを前記キャッシュメモリに書き込む手段と、前記専用のデータ転送路で相互に結合された他方のチャネル制御部に対して、前記専用のデータ転送路を介して前記書き込みデータを送信する手段と、前記他方のチャネル制御部から、前記他方のチャネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を前記専用のデータ転送路を介して受信する手段と、前記情報処理装置に対して、前記書き込みデータの前記キャッシュメモリへの書き込みが完了した旨の報告を送信する手段とを備えることを特徴とする。

【 0 0 2 6 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部であって、前記第2のキャッシュメモリに記憶されているデータの読み込みコマンドを前記第2のキャッシュメモリに送信する手段と、前記データを前記第2のキャッシュメモリから取得する手段と、前記取得したデータを前記第1のキャッシュメモリに書き込む手段と、前記専用のデータ転送路で相互に結合された他方のチャネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信す

る手段と、前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信する手段とを備えることを特徴とする。

【 0 0 2 7 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部であって、前記情報処理装置からアドレスが指定されたデータの読み出しコマンドを受信する手段と、前記指定されたアドレスのデータが前記第1のキャッシュメモリに記憶されているかどうかを判定する手段と、前記指定されたアドレスのデータが前記第1のキャッシュメモリに記憶されていない場合には、前記第2のキャッシュメモリへ前記データの読み込みコマンドを送信する手段と、前記第2のキャッシュメモリから前記データを取得する手段と、前記取得したデータを前記第1のキャッシュメモリに書き込む手段と、前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信する手段と、前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信する手段と、前記情報処理装置に対して前記データを送信する手段とを備えることを特徴とする。

【 0 0 2 8 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相

互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部であって、前記第 1 のキャッシュメモリに記憶されているデータの中から前記第 2 のキャッシュメモリへ書き出すデータを選定する手段と、前記第 2 のキャッシュメモリに前記選定したデータを書き込むための記憶領域を確保する手段と、前記選定したデータを前記第 2 のキャッシュメモリへ送信する手段とを備えることを特徴とする。

【 0 0 2 9 】

また本発明の一態様は、前記選定は最も長期間アクセスのなかったデータを選出することにより行われることを特徴とする。

【 0 0 3 0 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第 1 のキャッシュメモリと、相互に同一のデータを記憶する 2 つの第 2 のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第 2 のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置であって、前記記憶制御装置への電力供給が停止した場合に、継続して電力供給を行うためのバッテリーと、前記記憶制御装置への電力供給が停止した場合に、前記バッテリーに電力供給源を切り替える手段と、前記第 1 のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記第 2 のキャッシュメモリに未書き込みのものを前記第 2 のキャッシュメモリへ書き込む手段と、前記第 2 のキャッシュメモリへの書き込み完了後に、前記バッテリーから前記チャンネル制御部への供給電力を低減する手段と、前記第 2 のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記記憶装置に未書き込みのものを前記記

憶装置へ書き込む手段と、前記記憶装置への書き込み完了後に、前記バッテリから前記記憶制御装置への電力供給を遮断する手段とを備えることを特徴とする。

【 0 0 3 1 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されているキャッシュメモリと、前記チャネル制御部と前記ディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部に、前記情報処理装置から書き込みデータを受信するステップと、前記書き込みデータを前記キャッシュメモリに書き込むステップと、前記専用のデータ転送路で相互に結合された他方のチャネル制御部に対して、前記専用のデータ転送路を介して前記書き込みデータを送信するステップと、前記他方のチャネル制御部から、前記他方のチャネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を前記専用のデータ転送路を介して受信するステップと、前記情報処理装置に対して、前記書き込みデータの前記キャッシュメモリへの書き込みが完了した旨の報告を送信するステップとを実行させるためのプログラムに関する。

【 0 0 3 2 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャネル制御部に、前記第2のキャッシュメモリに記憶されているデータの読み込みコマンドを前記第2のキャッシュメモリに送信するステップと、前記データを前記第2のキャッシュ

メモリから取得するステップと、前記取得したデータを前記第1のキャッシュメモリに書き込むステップと、前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信するステップと、前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップとを実行させるためのプログラムに関する。

【 0 0 3 3 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部に、前記情報処理装置からアドレスが指定されたデータの読み出しコマンドを受信するステップと、前記指定されたアドレスのデータが前記第1のキャッシュメモリに記憶されているかどうかを判定するステップと、前記指定されたアドレスのデータが前記第1のキャッシュメモリに記憶されていない場合には、前記第2のキャッシュメモリへ前記データの読み込みコマンドを送信するステップと、前記第2のキャッシュメモリから前記データを取得するステップと、前記取得したデータを前記第1のキャッシュメモリに書き込むステップと、前記専用のデータ転送路で相互に結合された他方のチャンネル制御部に対して、前記専用のデータ転送路により前記取得したデータを送信するステップと、前記他方のチャンネル制御部から、前記他方のチャンネル制御部が備えるキャッシュメモリへの前記送信したデータの書き込みが完了した旨の報告を受信するステップと、前記情報処理装置に対して前記データを送信するステップとを実行させるためのプログラムに関する。

【 0 0 3 4 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャ

ネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置における前記チャンネル制御部に、前記第1のキャッシュメモリに記憶されているデータの中から前記第2のキャッシュメモリへ書き出すデータを選定するステップと、前記第2のキャッシュメモリに前記選定したデータを書き込むための記憶領域を確保するステップと、前記選定したデータを前記第2のキャッシュメモリへ送信するステップとを実行させるためのプログラムに関する。

【 0 0 3 5 】

また本発明の一態様は、前記選定は最も長期間アクセスのなかったデータを選出することにより行われることを特徴とするプログラムに関する。

【 0 0 3 6 】

また本発明の一態様は、情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、前記情報処理装置と前記記憶装置との間で授受されるデータを一時的に記憶し、前記記憶したデータを相互に記憶するための専用のデータ転送路で相互に結合されている前記チャンネル制御部に配置されている第1のキャッシュメモリと、相互に同一のデータを記憶する2つの第2のキャッシュメモリと、前記チャンネル制御部と前記ディスク制御部と前記第2のキャッシュメモリとを相互に接続する内部接続部とを備える記憶制御装置に、前記記憶制御装置への電力供給が停止した場合に継続して電力供給を行うためのバッテリーに電力供給源を切り替えるステップと、前記第1のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記第2のキャッシュメモリに未書き込みのものを前記第2のキャッシュメモリへ書き込むステップと、前記第2のキャッシュメモリへの書き込み完了後に、前記バッテリーから前記チャンネル制御部への供

給電力を低減するステップと、前記第2のキャッシュメモリに記憶されている、前記情報処理装置により更新されたデータであって、前記記憶装置に未書き込みのものを前記記憶装置へ書き込むステップと、前記記憶装置への書き込み完了後に、前記バッテリーから前記記憶制御装置への電力供給を遮断するステップとを実行させるためのプログラムに関する。

【0037】

【実施例】

<第1の実施例>

第1の実施例に係る計算機システムの全体構成を示すブロック図を図1に示す。

ディスク制御装置（記憶制御装置）100は4台のホスト計算機（情報処理装置）200と接続され、データのリード／ライトを受け付ける。各ホスト計算機200とディスク制御装置100間のアクセスルートは2重化されており、同一のホスト計算機200からは別のチャンネル制御部300に接続されるように構成される。こうすることにより片方のアクセスルートに障害が発生しても他方のアクセスルートによりディスクアクセス処理を継続することが可能となる。

【0038】

チャンネル制御部300は、ホストインタフェース制御部（ホストIF制御部）310を備え、ホスト計算機200とのインタフェース機能を有する。またキャッシュ制御部320及びローカルキャッシュ（第1のキャッシュメモリ）330を備え、記憶装置800から読み出したデータを一時的に格納しておくことによりディスクアクセスの高速化を図っている。キャッシュ制御部320は、ペアとなる別のチャンネル制御部300のキャッシュ制御部320とペア間接続部350を介して相互に接続されており、お互いのローカルキャッシュ330のデータを記憶し合うことによりデータの2重化を行っている。

チャンネル制御部300は内部IF制御部340を備え、内部接続部500を介してグローバルキャッシュ（第2のキャッシュメモリ）600、及びディスク制御部400に接続されている。

【0039】

ディスク制御部 4 0 0 は記憶装置 8 0 0 と接続するためのインタフェース機能を備える。ディスク制御部 4 0 0 が記憶装置 8 0 0 から読み出したデータは内部接続部 5 0 0 を介してグローバルキャッシュ 6 0 0 のデータ領域 6 0 1 に格納された後、チャンネル制御部 3 0 0 の内部 I F 制御部 3 4 0 及びキャッシュ制御部 3 2 0 を経由してローカルキャッシュ 3 3 0 のデータ領域 3 3 1 に格納される。

グローバルキャッシュ 6 0 0 は 2 つ設けられており、お互いの記憶するデータを相互に記憶し合うことによりデータの 2 重化を図っている。

【 0 0 4 0 】

また、ディスク制御装置 1 0 0 はバッテリー 7 0 0 により緊急時の電源を確保している。これにより停電によりディスク制御装置 1 0 0 への電力の供給が途絶えても、バッテリー 7 0 0 からの電力供給を受けている間にローカルキャッシュ 3 3 0 及びグローバルキャッシュ 6 0 0 に記憶されているデータを記憶装置 8 0 0 へ書き込むことができ、ローカルキャッシュ 3 3 0 及びグローバルキャッシュ 6 0 0 に記憶されているデータの消失を防止している。

【 0 0 4 1 】

図 2 は、図 1 に示す計算機システムの別の形態に係る計算機システムの全体構成を示すブロック図である。

この形態は図 1 におけるチャンネル制御部 3 0 0 とディスク制御部 4 0 0 が統合されたものである。制御部 9 0 0 はホスト計算機 2 0 0 とのインタフェースを司るホスト I F 制御部 9 1 0 と、記憶装置 8 0 0 とのインタフェースを司るディスク I F 制御部 9 6 0 を備えている。

【 0 0 4 2 】

この形態によれば、チャンネル制御部、ディスク制御部、及びキャッシュが同一パッケージ上で構成されるので、パッケージを追加していくことにより容易にシステムを拡張することが可能である。パッケージとは複数の機能をモジュール化して 1 つの部品として構成したものである。部品の交換等の保守・管理はパッケージ単位に行われる。図 1 の構成はシステムの最大規模を想定して最大規模でのコスト低減効果が最大になるようにしたものであり、大規模構成向きである。従って既にビジネスが安定している顧客が比較的大規模な構成を実現するときには有

効である。これに対して図 2 の構成は初期導入時点でのコスト効果が最大になるようにしたものであり、なおかつ拡張性を維持しているため小中規模から大規模構成向きである。従って、これからビジネスを開始する顧客やビジネス環境の変化が激しい顧客が状況に応じてシステム規模を変更できる柔軟性のあるシステムを実現するときには有効である。

【 0 0 4 3 】

またディスク I F 制御部 9 6 0 が記憶装置 8 0 0 から読み出したデータはキャッシュ制御部 9 2 0 を経由してローカルキャッシュ 9 3 0 のデータ領域 9 3 1 に格納される。内部接続部 5 0 0 やグローバルキャッシュ 6 0 0 を経由しないので、高速なデータの読み出しが可能である。

その他の部分の構成は図 1 に示した計算機システムの構成と同様であるため、以下では、図 1 に示した実施例に基づいて詳細に説明を行うこととする。

【 0 0 4 4 】

ホスト I F 制御部 3 1 0 及びキャッシュ制御部 3 2 0 の構成を示すブロック図を図 3 に示す。

ホスト I F 制御部 3 1 0 は、プロセッサ 3 1 1、メモリ 3 1 2、ホスト I F 回路 3 1 4、及び内部接続 I F 回路 3 1 5 を備えている。

【 0 0 4 5 】

プロセッサ 3 1 1 はメモリ 3 1 2 に記憶されている制御プログラム 3 1 3 を実行することにより、ホスト計算機 2 0 0 からのデータのリードライト要求を受け付け、ディスク制御装置 1 0 0 内の制御を行う。ホスト I F 回路 3 1 4 はホスト計算機 2 0 0 と接続され、データの授受を行うための回路を構成する。内部接続 I F 回路 3 1 5 はキャッシュ制御部 3 2 0 との接続のための回路を構成する。

【 0 0 4 6 】

キャッシュ制御部 3 2 0 は、キャッシュ制御部 I F 回路 3 2 1、バッファメモリ 3 2 2、内部接続 I F 回路 3 2 3、3 2 4、及びペア間接続 I F 回路 3 2 5 を備えている。

【 0 0 4 7 】

キャッシュ制御部 I F 回路 3 2 1 は、ローカルキャッシュ 3 3 0 と接続するた

めの回路を構成し、ローカルキャッシュ 3 3 0 との間のデータの授受を制御する。バッファメモリ 3 2 2 は、ローカルキャッシュ 3 3 0 との間のデータ授受の際に一時的にデータを格納するために用いられる。内部接続 I F 回路 3 2 3 は、ホスト I F 制御部と接続するための回路を構成する。内部接続 I F 回路 3 2 4 は、内部 I F 制御部 3 4 0 と接続するための回路を構成する。

【 0 0 4 8 】

ペア間接続 I F 回路 3 2 5 は、ペアとなる相手のチャンネル制御部 3 0 0 のキャッシュ制御部 3 2 0 と接続される。接続の様子を図 4 に示す。

ペアとなるチャンネル制御部 3 0 0 はお互いのローカルキャッシュ 3 3 0 のデータを共有することによりデータの 2 重化を行っている。データの 2 重化を行うためのコマンドやデータはペア間接続 I F 回路 3 2 5 を介して相手側のチャンネル制御部 3 0 0 に送られる。相互のペア間接続 I F 回路 3 2 5 の間はペア間接続部 3 5 0 により直結されている。ペア間接続部 3 5 0 は相互のローカルキャッシュ 3 3 0 のデータを 2 重化するために専用に使われた通信路である。そのため他の処理に影響を受けることなく相手側のローカルキャッシュ 3 3 0 とのデータの通信が行えるため、高速にデータの 2 重化を行うことができる。

【 0 0 4 9 】

またこの 2 重化はペアを組んでいるチャンネル制御部の間で行われるため、記憶制御装置の運転に影響を与えずに、チャンネル制御部に対する保守作業を行うことができる。例えば故障したチャンネル制御部を交換する場合を考える。もし、ペアを組まずに自己のチャンネル制御部内でデータを 2 重化している場合には、当該チャンネル制御部のキャッシュメモリに記憶されているデータを消失させないために、交換前にキャッシュメモリに記憶されているデータを当該チャンネル制御部以外の他の記憶領域に待避させる必要がある。しかし本実施例の場合にはペアを組むチャンネル制御部のキャッシュメモリにもデータが記憶されているため、データの待避は不要となる。これにより、24 時間 3 6 5 日の連続運転が要求される記憶制御装置の運転を停止させたり機能を低下させたりすることなく、保守・管理作業を行うことが可能となる。

【 0 0 5 0 】

なお、内部接続 I F 回路 3 1 5、内部接続 I F 回路 3 2 3、3 2 4、及びペア間接続 I F 回路 3 2 5 の回路構成は、同種、異種、もしくは同種・異種の混在のいずれの態様とすることもできる。

【 0 0 5 1 】

次に、ローカルキャッシュ 3 3 0 の構成を示すブロック図を図 5 に示す。

ローカルキャッシュ 3 3 0 は、データ領域 3 3 1 と制御領域 3 3 2 を備える。データ領域 3 3 1 にはホスト計算機 2 0 0 により記憶装置 8 0 0 から読み出されたデータが一時的に格納される。制御領域 3 3 2 にはデータ領域 3 3 1 に格納されているデータを管理するためのテーブルが記憶される。データ領域 3 3 1 に格納されているデータブロック毎に、"V a l i d"、"D i r t y"、"A d d r e s s"、"L o c k"、"O w n e r"、"P o i n t e r" の欄を有するテーブルを設けることによりデータを管理している。

【 0 0 5 2 】

なおデータ領域 3 3 1 に格納されるデータブロックは、どのような単位で格納されるようにすることも可能である。記憶装置 8 0 0 のブロック単位やシリンダの単位、あるいはトラックの単位等に限定されるものではない。またデータブロックのサイズは可変長とすることもできるし、固定長とすることもできる。

【 0 0 5 3 】

"V a l i d" 欄は、当該データブロックのデータが有効か否かを示す。ホスト計算機 2 0 0 からデータの読み出し要求があった場合に、当該データをデータ領域 3 3 1 に見つけることができても、当該データが有効でなければキャッシュアクセスはミスヒットとなる。

【 0 0 5 4 】

"D i r t y" 欄は、ローカルキャッシュ 3 3 0 に読み出されたデータがホスト計算機 2 0 0 により書き換えられているか否かを示す。書き換えられている場合は、当該データを記憶装置 8 0 0 へ書き戻しておく必要がある。書き換えられていなければ、当該データを記憶装置 8 0 0 へ書き戻す必要はない。

【 0 0 5 5 】

"A d d r e s s" 欄は、ローカルキャッシュ 3 3 0 に記憶されるデータの記憶

位置を示す。ホスト計算機 2 0 0 は、データを読み出す際に、データを論理アドレスで指定する。そのため、“A d d r e s s”欄には論理アドレスが記載される。しかし、本実施例におけるディスク制御装置 1 0 0 ではグローバルキャッシュ 6 0 0 は物理アドレスで管理されるため、制御領域 3 3 2 にはマッピングテーブルが設けられる。マッピングテーブルを参照して論理アドレスと物理アドレスの変換を行うことにより、ローカルキャッシュ 3 3 0 とグローバルキャッシュ 6 0 0 間でのデータ転送が可能となる。

【 0 0 5 6 】

“L o c k”欄は、ペアのローカルキャッシュ 3 3 0 間で相互に記憶されている当該データに対する処理を禁止することを表す。ペアとなるローカルキャッシュ 3 3 0 間のデータは、データ 2 重化専用の通信路であるペア間接続部 3 5 0 を経由して高速に 2 重化されるものの、完全同時に 2 重化することはできない。そのため極めて短時間ではあるがペア間でデータの不一致が生じる。ペア間でデータが不一致の間に、例えば片方のローカルキャッシュから当該データがリブレース（グローバルキャッシュ 6 0 0 への書き戻し）されてしまうと、誤ったデータがグローバルキャッシュ 6 0 0 や記憶装置 8 0 0 に記憶されることも起こりうる。このような問題を発生させないために“L o c k”欄が設けられ、L o c k が有効な間は当該データに対する操作は禁止される。

【 0 0 5 7 】

“O w n e r”欄は、ペアとなるローカルキャッシュ 3 3 0 のどちらが当該データを所有しているのかを示す。ペア間では相互にデータを 2 重化して記憶し合っているため、どちらのデータであるのかを管理するために“O w n e r”欄が設けられている。

【 0 0 5 8 】

“P o i n t e r”欄は、データ領域 3 3 1 に記憶されるデータと制御領域 3 3 2 に記憶される管理情報の対応付けを管理するための欄である。

【 0 0 5 9 】

次に、グローバルキャッシュ 6 0 0 の構成を示すブロック図を図 6 に示す。

基本的な構成はローカルキャッシュ 3 3 0 の構成と同一であるが、“L o c k”

欄と"Owner"欄の表す意味が異なる。

【0060】

"Lock"欄は、グローバルキャッシュ600上の当該データがローカルキャッシュ330に読み出されており、ホスト計算機200により更新される可能性があるため、他のローカルキャッシュ330への読み出しは禁止されている状態であることを表す。複数のローカルキャッシュ330にデータの読み出しを許してしまうと、それぞれ独立にホスト計算機200によって更新される可能性があり、データの一致性が保証できなくなるためである。

"Owner"欄は、当該データを読み出し中のローカルキャッシュ330を表す。

グローバルキャッシュ600は内部接続部500に接続されており、2つのグローバルキャッシュ600がペアとなってデータの2重化を行っている。グローバルキャッシュ600間のデータの2重化は、内部接続部500を介して相互にデータを転送することにより実現される。

【0061】

次に、内部接続部500の構成を示すブロック図を図7に示す。

図7には4入力4出力の例を示したが本実施例に係る内部接続部500は8入力8出力となる。しかし基本的な構成は変わらないので説明の簡単化のため4入力4出力の例で説明する。

【0062】

内部接続部500は、受信部510、送信部520、制御部530を備えている。受信部510は、内部接続部500に入力されてきたデータを適宜バッファ511に蓄えつつ、制御部530からの指令に従って、指定された送信部520のバッファ521へデータを転送する。送信部520はバッファ521に格納されたデータを順次出力する。図7では内部接続部500としてクロスバスイッチの構成をとった場合を示したが、クロスバスイッチの構成に限られることはなく、様々な構成をとることが可能である。例えば受信部510と送信部520の間を多段のスイッチ回路で接続するようにすることもできる。

【0063】

2つのグローバルキャッシュ600の間で2重化されるデータは内部接続部500を経由するので、転送経路が多い上、経路上での競合も起こりうる。そのためグローバルキャッシュ600の2重化には比較的長時間を要する。一方、ローカルキャッシュ330の間で2重化されるデータは、ペア間接続部350で直接に接続されているので、転送経路も短くまた経路上の競合もないため、短い時間でデータの2重化を行うことができる。

【0064】

次に、ホスト計算機200から本実施例に係るディスク制御装置100に対してデータアクセス要求があった場合の処理の流れを示すフローチャートを図8に示す。

ホスト計算機200からチャンネル制御部300にデータのアクセス要求があると（ステップS100）、ホストIF制御部310のプロセッサ311はアクセス要求の解析を行う。解析によりアクセスの種類（リード要求、ライト要求）やアクセスするデータのアドレス等を判別する。

【0065】

続いてプロセッサ311は、アクセスの種類に応じて、図9に示すコマンドをキャッシュ制御部320のキャッシュ制御部IF回路321に送信する。リード要求の場合は図9（d）のコマンドが送信され、ライト要求の場合は図9の（a）のコマンドと（b）のコマンド（データ）が送信される。

【0066】

キャッシュ制御部IF回路321はプロセッサ311から送信されたコマンドに従い、ローカルキャッシュ330の制御領域332に記録されている管理情報を検索して、コマンドに指定されたアドレスのデータがローカルキャッシュ330に記憶されているかどうかを確認する（ステップS101）。

【0067】

当該データがローカルキャッシュ330にある場合は、リード要求の場合であれば、キャッシュ制御部IF回路321が当該データをローカルキャッシュ330から読み出してホスト計算機200へ送信する（ステップS106、ステップS107）。ローカルキャッシュ330から読み出し完了の報告（ACK）を受

けると、キャッシュ制御部 I F 回路 3 2 1 はプロセッサ 3 1 1 に対してステータスを送信する。送信されるステータスは図 9 の (f) で示されるコマンドである。最後にプロセッサ 3 1 1 はホスト計算機 2 0 0 にデータの読み出し完了報告を行って (ステップ S 1 0 9) 処理を終了する。

【 0 0 6 8 】

なお本実施例においては、ローカルキャッシュ 3 3 0 の制御領域 3 3 2 の管理情報の検索や、ローカルキャッシュ 3 3 0 からのデータの読み出し等の制御は、キャッシュ制御部 I F 回路 3 2 1 が行う場合を例に説明したが、プロセッサ 3 1 1 が行う態様とすることも可能である。またこの場合、プロセッサ 3 3 1 は制御領域 3 3 2 の管理情報をメモリ 3 1 2 にコピーして、ローカルキャッシュ 3 3 0 の管理情報の検索を高速化するようにすることもできる。

【 0 0 6 9 】

また詳細は後述するが、グローバルキャッシュ 6 0 0 へのデータアクセス制御についても、キャッシュ制御部 I F 回路 3 2 1 が行う態様に限らず、プロセッサ 3 1 1 が行う態様とすることも可能である。

【 0 0 7 0 】

以上のリード要求の処理をフローチャートで表したものを図 1 5 に示す。図 1 5 に示すように、ローカルキャッシュ 3 3 0 は 2 つでペアを組んでいるが、データの読み出し時には相手のローカルキャッシュ 3 3 0 へのアクセスは行われない。

【 0 0 7 1 】

一方、ホスト計算機 2 0 0 からのアクセス要求がライト要求の場合は、キャッシュ制御部 I F 回路 3 2 1 は、ホスト計算機 2 0 0 から送信されバッファメモリ 3 2 2 に格納されている書き込みデータをローカルキャッシュ 3 3 0 に書き込む (ステップ S 1 0 8)。

【 0 0 7 2 】

ローカルキャッシュ 3 3 0 への書き込み処理の詳細は図 1 4 に示される。すなわち、まずキャッシュ制御部 I F 回路 3 2 1 は、ペアとなっている相手側のキャッシュ制御部 I F 回路 3 2 1 に対してローカルキャッシュ 3 3 0 をロックするよ

うに要求を出す。この要求は図 1 1 の (a) に示すコマンドを、ペア間接続部 3 5 0 を介して送信することにより行われる。相手からロック確保の応答 (図 1 1 の (b) に示すコマンド) を受け取り、自他ローカルキャッシュ 3 3 0 のロックを確保すると (ステップ S 5 0 0)、キャッシュ制御部 I F 回路 3 2 1 は、バッファメモリ 3 2 2 に格納されている書き込みデータを、ペア間接続部 3 5 0 を介して相手側のバッファメモリ 3 5 0 に送信する。そして相手側のキャッシュ制御部 I F 回路 3 2 1 により相手側のローカルキャッシュ 3 3 0 に書き込みが行われる (ステップ S 5 0 1)。続いて自分側のローカルキャッシュ 3 3 0 にデータの書き込みを行う (ステップ S 5 0 2)。なお、データをローカルキャッシュ 3 3 0 に書き込む際には制御領域 3 2 2 の "D i r t y" 欄にチェックを入れる。相互のローカルキャッシュ 3 3 0 にデータの書き込みが完了すると、ロックを解除した後、ホスト計算機 2 0 0 へ完了報告を送信し、処理を終了する (ステップ S 5 0 3、ステップ S 1 0 9)。以上のライト要求の処理をフローチャートで表したものを図 1 6 に示す。

【 0 0 7 3 】

ホスト計算機 2 0 0 はディスク制御装置 1 0 0 からの完了報告を受けるまでは次の処理を行うことができないので、ディスク制御装置 1 0 0 がホスト計算機 2 0 0 に対していかに早く完了報告を返すかが性能向上のポイントとなる。本実施例に係るディスク制御装置 1 0 0 では、ペアとなる相手側のローカルキャッシュ 3 3 0 へのデータの書き込みを、専用の通信路であるペア間接続部 3 5 0 を介して行うことにより、他の処理に影響を受けることなく高速に行うことができる。

【 0 0 7 4 】

またローカルキャッシュ 3 3 0 間でのデータ 2 重化のためのデータ転送は内部接続部 5 0 0 を使用せずに行われるので、内部接続部 5 0 0 の使用率を抑制することができる。そのため内部接続部 5 0 0 でのデータ競合が減少し、ホスト計算機 2 0 0 へのデータアクセスレスポンスを高速化できる。

【 0 0 7 5 】

次に、ホスト計算機 2 0 0 からのアクセス要求を受けたが、ローカルキャッシュ 3 3 0 に当該データが無い場合、すなわちキャッシュミスヒットの場合の処理

について説明する。

この場合はグローバルキャッシュ 6 0 0 に当該データがあるかどうかを確認する（ステップ S 1 0 2）。まずキャッシュ制御部 I F 回路 3 2 1 はプロセッサ 3 1 1 から送信されたコマンドに指定された当該データの論理アドレスを元に、制御領域 3 3 2 のマッピングテーブルを参照して物理アドレスを得る。そして、内部接続部 5 0 0 を介してグローバルキャッシュ 6 0 0 にコマンドを送信して、グローバルキャッシュ 6 0 0 の制御領域 6 0 2 に記録されている管理情報を検索して、当該データがグローバルキャッシュ 6 0 0 に記憶されているかどうかを確認する。

【 0 0 7 6 】

当該データがグローバルキャッシュ 6 0 0 に無ければ、ディスク制御部 4 0 0 に対してコマンドを送信し、当該データを記憶装置 8 0 0 から読み出して、グローバルキャッシュ 6 0 0 に格納させる（ステップ S 1 0 3）。グローバルキャッシュ 6 0 0 に格納されたデータは、内部接続部 5 0 0 を介してもう 1 つのグローバルキャッシュ 6 0 0 へも送られ、データの 2 重化が行われる。

【 0 0 7 7 】

なおここで、記憶装置 8 0 0 からグローバルキャッシュ 6 0 0 へ読み出したデータをいち早くホスト計算機 2 0 0 へ届けるための処理を優先させ、グローバルキャッシュ 6 0 0 上でのデータの 2 重化を後回しにする態様とすることも可能である。かかるグローバルキャッシュ 6 0 0 上のデータは記憶装置 8 0 0 にも記憶されているので、グローバルキャッシュ 6 0 0 上で消失しても問題は生じない。当該データが更新された場合に 2 重化を行うようにすることでデータの信頼性は確保できる。

【 0 0 7 8 】

続いて当該データをグローバルキャッシュ 6 0 0 上でロックを掛ける（ステップ S 1 0 4）。すなわち、グローバルキャッシュ 6 0 0 上の当該データを他のローカルキャッシュ 3 3 0 から読み出されないようにする。その処理の流れを図 1 0 のフローチャートに示す。

当該データがすでに他のローカルキャッシュ 3 3 0 に読み出されており、ロッ

クが掛けられている場合には（ステップ S 2 0 0）、当該ローカルキャッシュ 3 3 0 に対してロックを解放するように要求する（ステップ S 2 0 1）。どのローカルキャッシュ 3 3 0 がロックを掛けているのかは、制御領域 6 0 2 の "O w n e r" 欄で知ることができる。ロックが解放されるのを待った後（ステップ S 2 0 2）、ロックを掛け、他のローカルキャッシュ 3 3 0 から読み出されないようにしてから処理を終了する（ステップ S 2 0 3）。他のローカルキャッシュ 3 3 0 にロックが掛けられていなければ、直ちにロックを掛けて処理を終了する（ステップ S 2 0 0、ステップ S 2 0 3）。

【 0 0 7 9 】

続いてグローバルキャッシュ 6 0 0 上に格納された当該データをローカルキャッシュ 3 3 0 へ読み出す処理を行う（ステップ S 1 0 5）。その処理の流れを図 1 2 のフローチャートに示す。

まずグローバルキャッシュ 6 0 0 からローカルキャッシュ 3 3 0 にデータを転送する前に、ローカルキャッシュ 3 3 0 上に当該データを書き込むための空き領域があるかどうかを調べる（ステップ S 3 0 0）。この処理は、制御領域 3 3 2 の "V a l i d" 欄を検索して無効なデータの総容量がグローバルキャッシュ 6 0 0 から転送されるデータの総容量よりも大きいかどうかをチェックすることにより行うことができる。

【 0 0 8 0 】

十分な空き容量がある場合には、まずキャッシュ制御部 I F 回路 3 2 1 は、ペアとなっている相手側のキャッシュ制御部 I F 回路 3 2 1 に対してローカルキャッシュ 3 3 0 をロックするように要求を出してロックを確保する（ステップ S 3 0 2）。次にグローバルキャッシュ 6 0 0 からデータをバッファメモリ 3 2 2 に格納し、ペア間接続部 3 5 0 を介して相手側のバッファメモリ 3 5 0 にデータを送信するとともに、自分のローカルキャッシュ 3 3 0 にもデータの書き込みを行う（ステップ S 3 0 3、ステップ S 3 0 4）。相互のローカルキャッシュ 3 3 0 にデータの書き込みが完了すると、ロックを解除して処理を終了する（ステップ S 3 0 5）。この後の処理はホスト計算機 2 0 0 からのアクセス要求に応じて、上述した通りに行われる（ステップ S 1 0 6 乃至ステップ S 1 0 9）。

【 0 0 8 1 】

なお、グローバルキャッシュ 6 0 0 からローカルキャッシュ 3 3 0 にデータを転送するための空き領域がない場合は、ローカルキャッシュ 3 3 0 上のいずれかのデータをグローバルキャッシュ 6 0 0 に書き戻すことにより空き領域を確保する処理が必要になる（ステップ S 3 0 1）。その処理の流れを図 1 3 のフローチャートに示す。

【 0 0 8 2 】

まずキャッシュ制御部 I F 回路 3 2 1 は、ペアとなっている相手側のキャッシュ制御部 I F 回路 3 2 1 に対してローカルキャッシュ 3 3 0 をロックするように要求を出してロックを確保する（ステップ S 4 0 0）。続いて所定のアルゴリズムにより特定したグローバルキャッシュ 6 0 0 へ書き出されるデータの D i r t y ビットを制御領域 3 3 2 により調べる（ステップ S 4 0 1）。所定のアルゴリズムとしては、最も長期間アクセスのなかったデータをキャッシュから書き出す L R U (Least Recently Used) 方式が一般的であるが、他のアルゴリズムとすることもできる。

【 0 0 8 3 】

D i r t y ビットがセットされていなければグローバルキャッシュ 6 0 0 へデータを書き出す必要はないが、D i r t y ビットがセットされている場合はグローバルキャッシュ 6 0 0 へデータを書き出す必要があるので、グローバルキャッシュ 6 0 0 に当該データを書き込むための空き領域があるかどうかを調べる（ステップ S 4 0 2）。グローバルキャッシュ 6 0 0 に十分な空き領域がない場合はグローバルキャッシュ 6 0 0 のデータを記憶装置 8 0 0 へ書き出して、領域を確保する（ステップ S 4 0 3）。

【 0 0 8 4 】

続いてグローバルキャッシュ 6 0 0 上の空き領域にローカルキャッシュ 3 3 0 からデータを書き出す（ステップ S 4 0 4）。書き出しは 2 つのグローバルキャッシュ 6 0 0 に対して行われる。グローバルキャッシュ 6 0 0 にデータが書き出された後は、もはや当該データは D i r t y ではないので、D i r t y ビットをリセットする（ステップ S 4 0 5）。続いて当該データが記憶されていたローカ

ルキャッシュ 3 3 0 上の領域を解放する必要がある場合には（ステップ S 4 0 6）、当該データの V a l i d ビットをリセットする（ステップ S 4 0 7）。

【 0 0 8 5 】

そしてペアとなっている相手のローカルキャッシュ 3 3 0 に対して当該データのグローバルキャッシュ 6 0 0 への書き出しが完了した旨の報告を行う（ステップ S 4 0 8）。この報告をうけた相手側のローカルキャッシュ 3 3 0 では、制御領域 3 3 2 の V a l i d ビットがリセットされる。最後にローカルキャッシュ 3 3 0 のロックを解放して（ステップ S 4 0 9）処理を終了する。

【 0 0 8 6 】

次に、チャンネル制御部に障害が発生したときの処理を示すフローチャートを図 1 7 に示す。

まずホスト I F 制御部 3 1 0 のプロセッサ 3 1 1 は、チャンネル制御部 3 0 0 に異常がないかを常時監視している（ステップ S 8 0 0）。プロセッサ 3 1 1 はチャンネル制御部 3 0 0 に異常を検出すると、それ以降ローカルキャッシュ 3 3 0 を使用しないようにし、ローカルキャッシュ 3 3 0 を無効化する（ステップ S 8 0 1）。続いて既にローカルキャッシュ 3 3 0 上に記憶されているデータをグローバルキャッシュ 6 0 0 上に書き出す（ステップ S 8 0 2）。このとき書き出されるデータは制御領域 3 3 2 上で D i r t y ビットがセットされているデータである。次に当該ローカルキャッシュ 3 3 0 上から書き出されたデータのロックをグローバルキャッシュ 6 0 0 上で解除する（ステップ S 8 0 3）。他のローカルキャッシュ 3 3 0 に読み出し可能とするためである。この状態でチャンネル制御部 3 0 0 が正常に復帰するまで待つが（ステップ S 8 0 4）、その間のホスト計算機 2 0 0 からのデータアクセスは、ペアとなっているもう一方のチャンネル制御部 3 0 0 により継続される。チャンネル制御部 3 0 0 が正常に復帰したら、ローカルキャッシュ 3 3 0 を有効にし、使用再開する（ステップ S 8 0 5）。

これにより、チャンネル制御部 3 0 0 に障害が発生しても、ペアを組んでいるもう片方のチャンネル制御部 3 0 0 を使用することにより、システムを停止させることなく業務を継続することができる。

【 0 0 8 7 】

次に本実施例に係るディスク制御装置 1 0 0 の緊急デスページ処理を示すフローチャートを図 1 8 に示す。

緊急デスページ処理とは、突然の停電等によりディスク制御装置 1 0 0 に対する電力の供給がストップした場合にキャッシュメモリ上に記憶されているデータを保護するための技術である。停電時には電力の供給はバッテリー 7 0 0 により行われるが、ディスク制御装置 1 0 0 の電力消費量は少なくないので、バッテリー 7 0 0 による電力供給はそれほど長時間行える訳ではない。そのため、緊急デスページ処理は極力短時間で完了することが要求される。

【 0 0 8 8 】

ディスク制御装置 1 0 0 は電力の供給がバッテリー 7 0 0 からの供給に切り替わったことを検出すると緊急デスページ処理を開始する。まず、ローカルキャッシュ 3 3 0 に記憶されているデータの中でグローバルキャッシュ 6 0 0 に未反映のデータがないかを制御領域 3 3 2 の D i r t y ビットを検索することにより調査する（ステップ S 9 0 0）。グローバルキャッシュ 6 0 0 に未反映のデータは 2 つのグローバルキャッシュ 6 0 0 に書き出される（ステップ S 9 0 1）。全ての D i r t y データをグローバルキャッシュ 6 0 0 に書き出す処理が完了したら（ステップ S 9 0 2）、チャネル制御部 3 0 0 の電源を遮断する（ステップ S 9 0 3）。電力消費量を少しでも少なくし、バッテリー 7 0 0 による電力供給をできるだけ長時間維持するためである。

【 0 0 8 9 】

次にグローバルキャッシュ 6 0 0 に記憶されているデータの中で記憶装置 8 0 0 に未反映のデータがないかを制御領域 6 0 2 の D i r t y ビットを検索することにより調査する（ステップ S 9 0 4）。記憶装置 8 0 0 に未反映のデータはディスク制御部 4 0 0 を経由して記憶装置 8 0 0 に書き出される（ステップ S 9 0 5）。全ての D i r t y データを記憶装置 8 0 0 に書き出す処理が完了したら（ステップ S 9 0 6）、グローバルキャッシュ 6 0 0、ディスク制御部 4 0 0 及び記憶装置 8 0 0 の電源を遮断して（ステップ S 9 0 7）、処理を終了する。

【 0 0 9 0 】

なおステップ S 9 0 3 において、チャネル制御部 3 0 0 全体への電力供給を遮

断する態様のみならず、チャンネル制御部 3 0 0 の一部の部位への電力供給を遮断し、チャンネル制御部 3 0 0 への供給電力を低減させる態様とすることもできる。例えば、チャンネル制御部 3 0 0 の中でも比較的電力消費の大きな部位であるプロセッサ 3 1 1 への電力供給を遮断し、供給電力を低減させる態様とすることも可能である。

【 0 0 9 1 】

このような処理を行うことにより、突然の停電等が発生してもキャッシュメモリ上のデータを消失させることがなく、システムの信頼性を向上させることが可能となる。

【 0 0 9 2 】

以上のようにして、ホスト計算機 2 0 0 からディスク制御装置 1 0 0 に対してデータアクセス要求があった場合の処理が行われる。

これによれば、ローカルキャッシュ 3 3 0 及びグローバルキャッシュ 6 0 0 に記憶されるデータは 2 重化されているのでデータの信頼性を向上させることができる。

その上ローカルキャッシュ 3 3 0 が各チャンネル制御部 3 0 0 に分散して配置されているので、各ホスト計算機 2 0 0 は他のホスト計算機 2 0 0 からのデータアクセス要求との競合なしにキャッシュメモリへのアクセスを行うことができ、高速なデータアクセスを実現している。

【 0 0 9 3 】

さらにデータの 2 重化のためにペアを組んでいるローカルキャッシュ間は専用のペア間接続部 3 5 0 により結合されているので、データの 2 重化処理のオーバーヘッドを減らすことができ、ホスト計算機 2 0 0 からのデータアクセスのさらなる高速化を実現している。またローカルキャッシュ 3 3 0 間でのデータ 2 重化のためのデータ転送は内部接続部 5 0 0 を使用せずに行われるので、内部接続部 5 0 0 の使用率を抑制することができる。そのため内部接続部 5 0 0 でのデータ競合が減少し、ホスト計算機 2 0 0 へのデータアクセスレスポンスを高速化できる。

【 0 0 9 4 】

また、ホスト計算機 2 0 0 から要求のあったデータがキャッシュメモリに存在しない場合には、従来であれば磁気ディスクまでデータを読みに行く必要があったが、本実施例に係るディスク制御装置 1 0 0 ではローカルキャッシュ 3 3 0 にデータが存在しない場合でもグローバルキャッシュ 6 0 0 に当該データが存在すれば、磁気ディスクまでデータを読みに行く必要はないため、ホスト計算機 2 0 0 からのデータアクセスをより高速に行うことが可能となっている。

【 0 0 9 5 】

また、ホスト計算機 2 0 0 からのアクセスデータがローカルキャッシュ 3 3 0 に記憶されていなかった場合でも、グローバルキャッシュ 6 0 0 に当該データが記憶されていれば記憶装置 8 0 0 にまでアクセスをする必要が無い。そのため、ホスト計算機 2 0 0 からのデータのアクセス速度を高速化することが可能となる。

【 0 0 9 6 】

< 第 2 の実施例 >

なお、第 1 の実施例に係るディスク制御装置 1 0 0 において、ペア間接続部 3 5 0 を設けない態様も考えられる。このような態様であってもペアとなる 2 つのローカルキャッシュ 3 3 0 間の通信を高速に行うことによりデータの 2 重化を高速に行うことができる。

すなわち、ペア間接続部 3 5 0 がない場合には 2 つのローカルキャッシュ 3 3 0 間のデータ通信は内部接続部 5 0 0 を経由して行われることになるが、図 7 に示す回路において、相互にペアとなっているチャネル制御部 3 0 0 に接続される受信部 5 1 0 から送信部 5 2 0 へのデータ転送を他のデータ転送に優先、または占有して行うことにより、2 つのペア間はあたかも専用通信路で結ばれているかのごとくデータ転送を行うことが可能となる。これによりペア間のデータ転送は、第 1 の実施例におけるディスク制御装置 1 0 0 と同様、他のデータ転送に影響を受けることなく高速に行うことができる。

【 0 0 9 7 】

< 第 3 の実施例 >

次に第 3 の実施例に係る計算機システムの全体構成を示すブロック図を図 1 9

に示す。

第 3 の実施例においては、第 1 の実施例に係るディスク制御装置 1 0 0 に対し、グローバルキャッシュ 6 0 0 が存在しない態様の例を示す。このような態様においては、あるローカルキャッシュ 3 3 0 に読み出されているデータを他のローカルキャッシュ 3 3 0 に読み出されないように管理するための制御メモリ 1 0 0 0 が内部接続部 5 0 0 に設けられる。

【 0 0 9 8 】

制御メモリ 1 0 0 0 の構成を図 2 0 に示す。

制御メモリ 1 0 0 0 は、ローカルキャッシュ 3 3 0 に読み出されているデータを管理するための制御領域 1 0 1 0 を有している。制御領域 1 0 1 0 は、"A d d r e s s"欄、"L o c k"欄、"O w n e r"欄から構成される。

"A d d r e s s"欄は、ローカルキャッシュ 3 3 0 に読み出されているデータの物理アドレス情報を示す。

"L o c k"欄は、"A d d r e s s"欄で特定されるデータがローカルキャッシュ 3 3 0 に読み出されており、ホスト計算機 2 0 0 により更新される可能性があるため、他のローカルキャッシュ 3 3 0 への読み出しは禁止されている状態であることを表す。複数のローカルキャッシュ 3 3 0 にデータの読み出しを許してしまうと、それぞれ独立にホスト計算機 2 0 0 によって更新される可能性があり、データの一致性が保証できなくなるためである。

"O w n e r"欄は、当該データを読み出し中のローカルキャッシュ 3 3 0 を表す。

第 3 の実施例に係るディスク制御装置 1 0 0 のその他の構成は、第 1 の実施例に係るディスク制御装置 1 0 0 と同様である。

【 0 0 9 9 】

次に、ホスト計算機 2 0 0 から本実施例に係るディスク制御装置 1 0 0 に対してデータアクセス要求があった場合の処理の流れを示すフローチャートを図 2 1 に示す。

ホスト計算機 2 0 0 からチャンネル制御部 3 0 0 にデータのアクセス要求があると（ステップ S 1 0 0 0）、ホスト I F 制御部 3 1 0 のプロセッサ 3 1 1 はアク

セス要求の解析を行う。解析によりアクセスの種類（リード要求、ライト要求）やアクセスするデータのアドレス等を判別する。

【 0 1 0 0 】

続いてプロセッサ 3 1 1 は、アクセスの種類に応じて、図 9 に示すコマンドをキャッシュ制御部 3 2 0 のキャッシュ制御部 I F 回路 3 2 1 に送信する。リード要求の場合は図 9（d）のコマンドが送信され、ライト要求の場合は図 9 の（a）のコマンドと（b）のコマンド（データ）が送信される。

【 0 1 0 1 】

キャッシュ制御部 I F 回路 3 2 1 はプロセッサ 3 1 1 から送信されたコマンドに従い、ローカルキャッシュ 3 3 0 の制御領域 3 3 2 に記録されている管理情報を検索して、コマンドに指定されたアドレスのデータがローカルキャッシュ 3 3 0 に記憶されているかどうかを確認する（ステップ S 1 0 0 1）。

【 0 1 0 2 】

当該データがローカルキャッシュ 3 3 0 にある場合は、リード要求の場合であれば、キャッシュ制御部 I F 回路 3 2 1 が当該データをローカルキャッシュ 3 3 0 から読み出してホスト計算機 2 0 0 へ送信する（ステップ S 1 0 0 4、ステップ S 1 0 0 5）。ローカルキャッシュ 3 3 0 から読み出し完了の報告（ACK）を受けると、キャッシュ制御部 I F 回路 3 2 1 はプロセッサ 3 1 1 に対してステータスを送信する。送信されるステータスは図 9 の（f）で示されるコマンドである。最後にプロセッサ 3 1 1 はホスト計算機 2 0 0 にデータの読み出し完了報告を行って（ステップ S 1 0 0 7）処理を終了する。

【 0 1 0 3 】

一方、ホスト計算機 2 0 0 からのアクセス要求がライト要求の場合は、キャッシュ制御部 I F 回路 3 2 1 は、ホスト計算機 2 0 0 から送信されバッファメモリ 3 2 2 に格納されている書き込みデータをローカルキャッシュ 3 3 0 に書き込む（ステップ S 1 0 0 6）。

【 0 1 0 4 】

ローカルキャッシュ 3 3 0 への書き込み処理の詳細は図 2 5 に示される。すなわち、まずキャッシュ制御部 I F 回路 3 2 1 は、ペアとなっている相手側のキャ

ッシュ制御部 I F 回路 3 2 1 に対してローカルキャッシュ 3 3 0 をロックするように要求を出す。この要求は図 1 1 の (a) に示すコマンドを、ペア間接続部 3 5 0 を介して送信することにより行われる。相手からロック確保の応答 (図 1 1 の (b) に示すコマンド) を受け取り、自他ローカルキャッシュ 3 3 0 のロックを確保すると (ステップ S 1 4 0 0) 、キャッシュ制御部 I F 回路 3 2 1 は、バッファメモリ 3 2 2 に格納されている書き込みデータを、ペア間接続部 3 5 0 を介して相手側のバッファメモリ 3 5 0 に送信する。そして相手側のキャッシュ制御部 I F 回路 3 2 1 により相手側のローカルキャッシュ 3 3 0 に書き込みが行われる (ステップ S 1 4 0 1) 。続いて自分側のローカルキャッシュ 3 3 0 にデータの書き込みを行う (ステップ S 1 4 0 2) 。なお、データをローカルキャッシュ 3 3 0 に書き込む際には制御領域 3 2 2 の " D i r t y " 欄にチェックを入れる。相互のローカルキャッシュ 3 3 0 にデータの書き込みが完了すると、ロックを解除した後、ホスト計算機 2 0 0 へ完了報告を送信し、処理を終了する (ステップ S 1 4 0 3 、ステップ S 1 0 0 7) 。

【 0 1 0 5 】

ホスト計算機 2 0 0 はディスク制御装置 1 0 0 からの完了報告を受けるまでは次の処理を行うことができないので、ディスク制御装置 1 0 0 は、ホスト計算機 2 0 0 に対していかに早く完了報告を返すかがポイントとなる。本実施例に係るディスク制御装置 1 0 0 では、ペアとなる相手側のローカルキャッシュ 3 3 0 へもデータの書き込みを、専用の通信路であるペア間接続部 3 5 0 を介して行うことにより、他の処理に影響を受けることなく高速に行うことができる。

【 0 1 0 6 】

次に、ホスト計算機 2 0 0 からのアクセス要求を受けたが、ローカルキャッシュ 3 3 0 に当該データが無い場合、すなわちキャッシュミスヒットの場合の処理について説明する。この場合は記憶装置 8 0 0 に記憶された当該データをローカルキャッシュ 3 3 0 へ読み出す処理を行う (ステップ S 1 0 0 2) 。その処理の流れを図 2 3 のフローチャートに示す。

【 0 1 0 7 】

まず記憶装置 8 0 0 からローカルキャッシュ 3 3 0 にデータを転送する前に、

ローカルキャッシュ 3 3 0 上に当該データを書き込むための空き領域があるかどうかを調べる（ステップ S 1 2 0 0）。この処理は、制御領域 3 3 2 の”V a l i d”欄を検索して無効なデータの総容量が記憶装置 8 0 0 から転送されるデータの総容量よりも大きいかどうかをチェックすることにより行うことができる。

【 0 1 0 8 】

十分な空き容量がある場合には、まずキャッシュ制御部 I F 回路 3 2 1 は、ペアとなっている相手側のキャッシュ制御部 I F 回路 3 2 1 に対してローカルキャッシュ 3 3 0 をロックするように要求を出してロックを確保する（ステップ S 1 2 0 2）。次に記憶装置 8 0 0 からデータをバッファメモリ 3 2 2 に格納し、ペア間接続部 3 5 0 を介して相手側のバッファメモリ 3 5 0 にデータを送信するとともに、自分のローカルキャッシュ 3 3 0 にもデータの書き込みを行う（ステップ S 1 2 0 3、ステップ S 1 2 0 4）。相互のローカルキャッシュ 3 3 0 にデータの書き込みが完了すると、ロックを解除して処理を終了する（ステップ S 1 2 0 5）。この後の処理はホスト計算機 2 0 0 からのアクセス要求に応じて、上述した通りに行われる（ステップ S 1 0 0 4 乃至ステップ S 1 0 0 7）。

【 0 1 0 9 】

なお、記憶装置 8 0 0 からローカルキャッシュ 3 3 0 にデータを転送するための空き領域がない場合は、ローカルキャッシュ 3 3 0 上のいずれかのデータを記憶装置 8 0 0 に書き戻すことにより空き領域を確保する処理が必要になる（ステップ S 1 2 0 1）。その処理の流れを図 2 4 のフローチャートに示す。

【 0 1 1 0 】

まずキャッシュ制御部 I F 回路 3 2 1 は、ペアとなっている相手側のキャッシュ制御部 I F 回路 3 2 1 に対してローカルキャッシュ 3 3 0 をロックするように要求を出してロックを確保する（ステップ S 1 3 0 0）。続いて所定のアルゴリズムにより特定した記憶装置 8 0 0 へ書き出されるデータの D i r t y ビットを制御領域 3 3 2 により調べる（ステップ S 1 3 0 1）。所定のアルゴリズムとしては、最も長期間アクセスのなかったデータをキャッシュから書き出す L R U 方式が一般的であるが、他のアルゴリズムとすることもできる。

【 0 1 1 1 】

D i r t yビットがセットされていなければ記憶装置800へデータを書き出す必要はないが、D i r t yビットがセットされている場合は記憶装置800にローカルキャッシュ330からデータを書き出す（ステップS1302）。書き出しはディスク制御部400を経由することにより行われる。記憶装置800にデータが書き出された後は、もはや当該データはD i r t yではないので、D i r t yビットをリセットする（ステップS1303）。続いて当該データが記憶されていたローカルキャッシュ330上の領域を解放する必要がある場合には（ステップS1304）、当該データのV a l i dビットをリセットする（ステップS1305）。

【0112】

そしてペアとなっている相手のローカルキャッシュ330に対して当該データの記憶装置800への書き出しが完了した旨の報告を行う（ステップS1306）。この報告をうけた相手側のローカルキャッシュ330では、制御領域332のV a l i dビットがリセットされる。最後にローカルキャッシュ330のロックを解放して（ステップS1307）処理を終了する。

【0113】

以上のようにしてローカルキャッシュ330に読み出したデータを制御メモリ1000でロックを掛ける（ステップS1003）。すなわち、ローカルキャッシュ330に読み出した当該データを他のローカルキャッシュ330から読み出されないようにする。その処理の流れを図22のフローチャートに示す。

【0114】

当該データがすでに他のローカルキャッシュ330に読み出されており、ロックが掛けられている場合には（ステップS1100）、当該ローカルキャッシュ330に対してロックを解放するように要求する（ステップS1101）。どのローカルキャッシュ330がロックを掛けているのかは、制御領域1010の”O w n e r”欄で知ることができる。ロックが解放されるのを待った後（ステップS1102）、ロックを掛け、他のローカルキャッシュ330から読み出されないようにしてから処理を終了する（ステップS1103）。他のローカルキャッシュ330にロックが掛けられていなければ、直ちにロックを掛けて処理を終

了する（ステップ S 1 1 0 0、ステップ S 1 1 0 3）。

この後の処理はホスト計算機 2 0 0 からのアクセス要求に応じて、上述した通りに行われる（ステップ S 1 0 0 4 乃至ステップ S 1 0 0 7）。

【 0 1 1 5 】

以上のようにして、ホスト計算機 2 0 0 からディスク制御装置 1 0 0 に対してデータアクセス要求があった場合の処理が行われる。

これによれば、ローカルキャッシュ 3 3 0 に記憶されるデータは 2 重化されているのでデータの信頼性を向上させることができる。

その上ローカルキャッシュ 3 3 0 が各チャネル制御部 3 0 0 に分散して配置されているので、各ホスト計算機 2 0 0 は他のホスト計算機 2 0 0 からのデータアクセス要求との競合なしにキャッシュメモリへのアクセスを行うことができ、高速なデータアクセスを実現している。

【 0 1 1 6 】

さらにデータの 2 重化のためにペアを組んでいるローカルキャッシュ間には専用のペア間接続部 3 5 0 により結合されているので、データの 2 重化処理のオーバーヘッドを減らすことができ、ホスト計算機 2 0 0 からのデータアクセスのさらなる高速化を実現している。またローカルキャッシュ 3 3 0 間でのデータ 2 重化のためのデータ転送は内部接続部 5 0 0 を使用せずに行われるので、内部接続部 5 0 0 の使用率を抑制することができる。そのため内部接続部 5 0 0 でのデータ競合が減少し、ホスト計算機 2 0 0 へのデータアクセスレスポンスを高速化できる。

【 0 1 1 7 】

以上本実施の形態に係る実施例について説明したが、上記の実施例は本発明の理解を容易にするためのものであり、本発明を限定して解釈するためのものではない。本発明はその趣旨を逸脱することなく変更、改良され得ると共に、本発明にはその等価物も含まれる。

【 0 1 1 8 】

【発明の効果】

記憶制御装置のキャッシュメモリに記憶されるデータのアクセス高速化を図る

ことができる。

【図面の簡単な説明】

【図 1】 第 1 の実施例に係る計算機システムの全体構成を示すブロック図である。

【図 2】 第 1 の実施例の別の形態に係る計算機システムの全体構成を示すブロック図である。

【図 3】 第 1 の実施例に係るホスト I F 制御部及びキャッシュ制御部の構成を示すブロック図である。

【図 4】 第 1 の実施例に係るチャンネル制御部の構成を示すブロック図である。

【図 5】 第 1 の実施例に係るローカルキャッシュメモリの構成を示すブロック図である。

【図 6】 第 1 の実施例に係るグローバルキャッシュメモリの構成を示すブロック図である。

【図 7】 第 1 の実施例に係る内部接続部の構成を示すブロック図である。

【図 8】 第 1 の実施例に係るデータアクセス処理を示すフローチャートである。

【図 9】 第 1 の実施例に係るデータアクセス処理を行うためのコマンドを示す図である。

【図 1 0】 第 1 の実施例に係るグローバルキャッシュメモリのロック確保処理を示すフローチャートである。

【図 1 1】 第 1 の実施例に係るデータアクセス処理を行うためのコマンドを示す図である。

【図 1 2】 第 1 の実施例に係るローカルキャッシュメモリのステージング処理を示すフローチャートである。

【図 1 3】 第 1 の実施例に係るローカルキャッシュメモリのデステージング処理を示すフローチャートである。

【図 1 4】 第 1 の実施例に係るローカルキャッシュメモリへのデータ書き込み処理を示すフローチャートである。

【図 1 5】 第 1 の実施例に係るローカルキャッシュメモリからのデータ読み出し処理を示すフローチャートである。

【図 1 6】 第 1 の実施例に係るローカルキャッシュメモリへのデータ書き込み処理を示すフローチャートである。

【図 1 7】 第 1 の実施例に係るチャネル制御部に障害が発生したときの処理を示すフローチャートである。

【図 1 8】 第 1 の実施例に係る記憶制御装置における緊急デステージ処理を示すフローチャートである。

【図 1 9】 第 3 の実施例に係る計算機システムの全体構成を示すブロック図である。

【図 2 0】 第 3 の実施例に係る制御メモリの構成を示すブロック図である。

【図 2 1】 第 3 の実施例に係るデータアクセス処理を示すフローチャートである。

【図 2 2】 第 3 の実施例に係るロック確保処理を示すフローチャートである。

【図 2 3】 第 3 の実施例に係るローカルキャッシュメモリのステージング処理を示すフローチャートである。

【図 2 4】 第 3 の実施例に係るローカルキャッシュメモリのデステージング処理を示すフローチャートである。

【図 2 5】 第 3 の実施例に係るローカルキャッシュメモリへのデータ書き込み処理を示すフローチャートである。

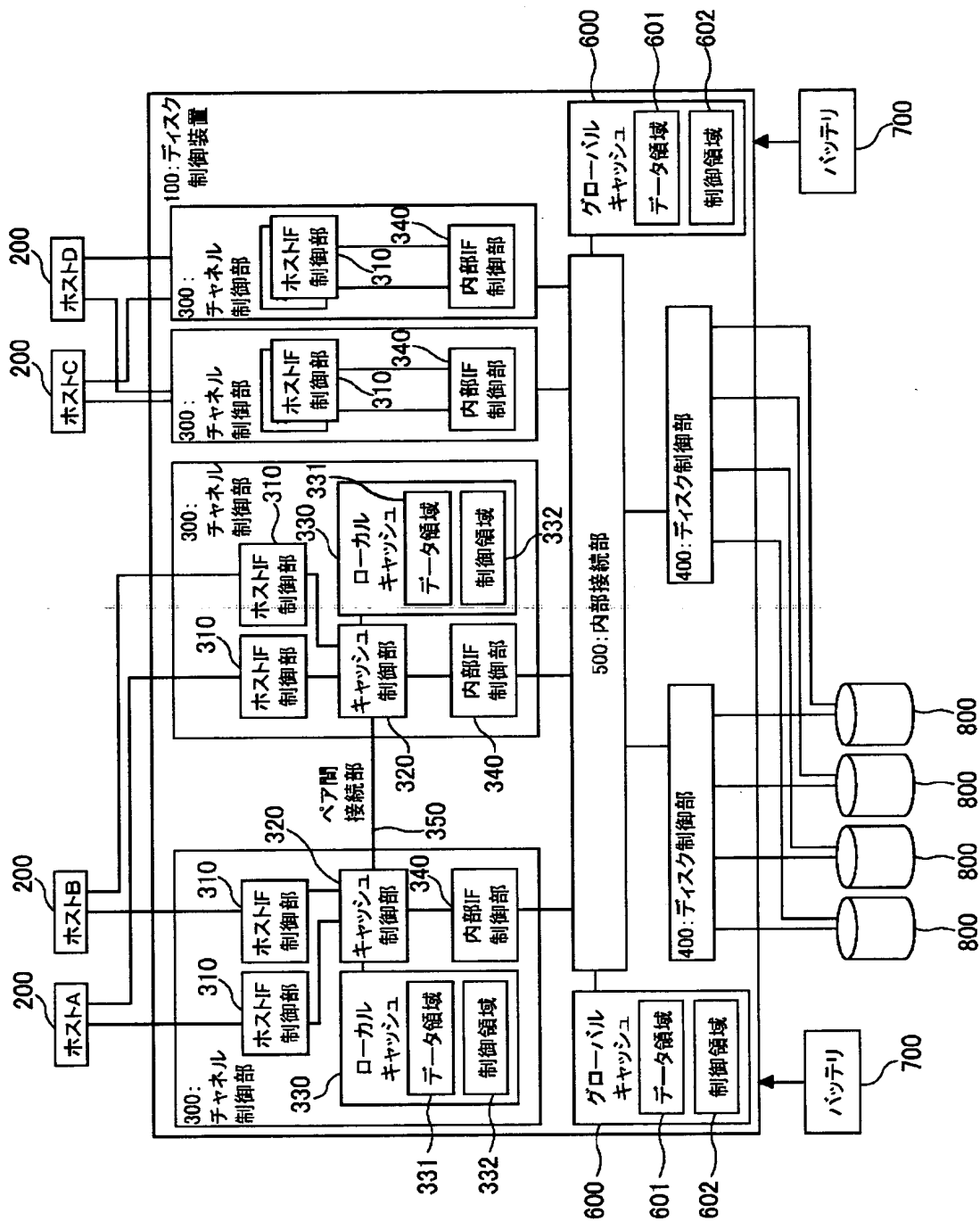
【符号の説明】

1 0 0	ディスク制御装置	2 0 0	ホスト計算機
3 0 0	チャネル制御部	3 1 0	ホスト I F 制御部
3 1 1	プロセッサ	3 1 2	メモリ
3 1 3	制御プログラム	3 1 4	ホスト I F 回路
3 1 5	内部接続 I F 回路	3 2 0	キャッシュ制御部
3 2 1	キャッシュ制御部 I F 回路	3 2 2	バッファメモリ

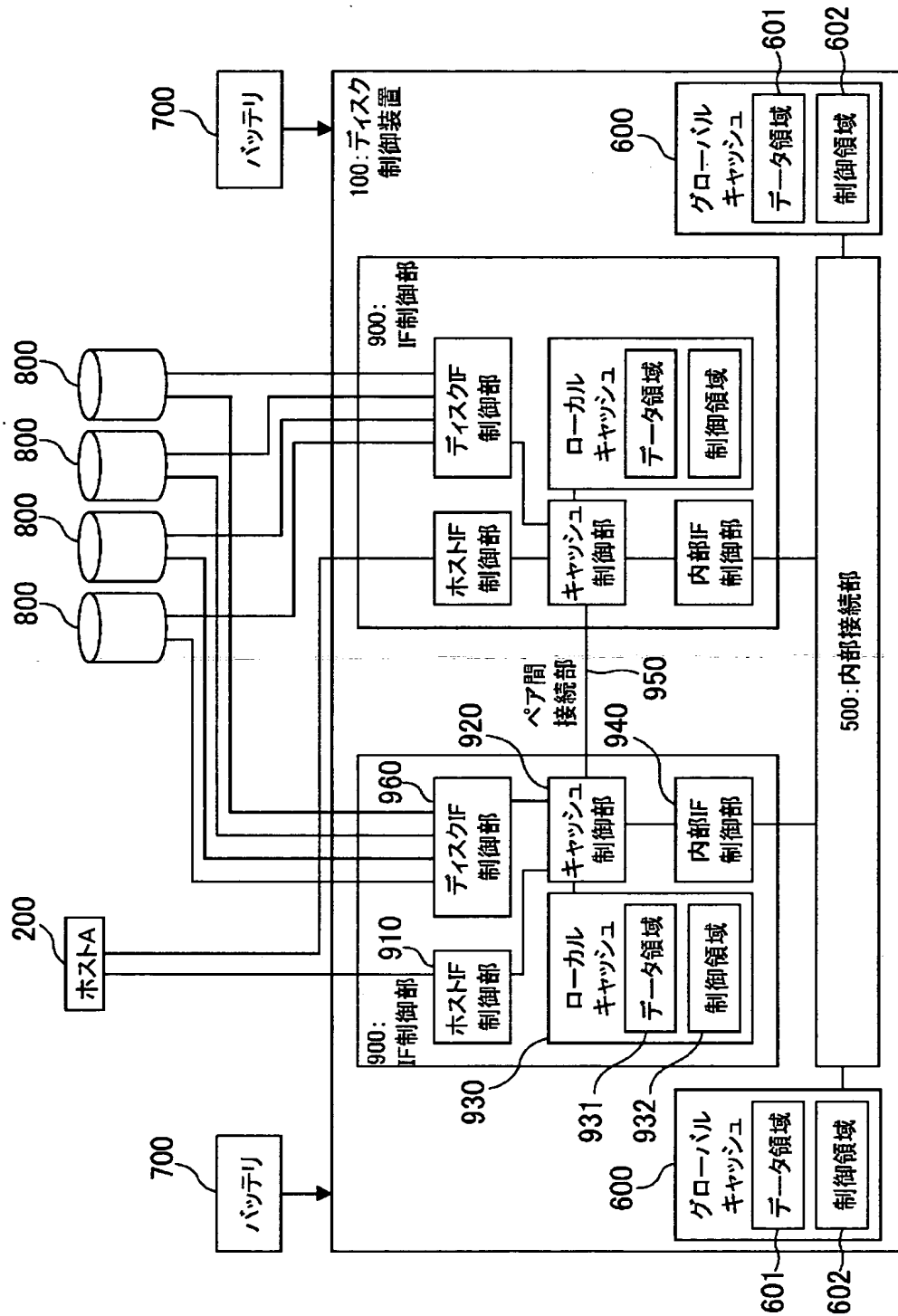
3 2 3	内部接続 I F 回路	3 2 4	内部接続 I F 回路
3 2 5	ペア間接続 I F 回路	3 3 0	ローカルキャッシュ
3 3 1	データ領域	3 3 2	制御領域
3 4 0	内部 I F 制御部	3 5 0	ペア間接続部
4 0 0	ディスク制御部	5 0 0	内部接続部
5 1 0	受信部	5 1 1	バッファ
5 2 0	送信部	5 2 1	バッファ
5 3 0	制御部	6 0 0	グローバルキャッシュ
6 0 1	データ領域	6 0 2	制御領域
7 0 0	バッテリー	8 0 0	記憶装置
9 0 0	I F 制御部	9 1 0	ホスト I F 制御部
9 2 0	キャッシュ制御部	9 3 0	ローカルキャッシュ
9 3 1	データ領域	9 3 2	制御領域
9 4 0	内部 I F 制御部	9 5 0	ペア間接続部
9 6 0	ディスク I F 制御部	1 0 0 0	制御メモリ
1 0 1 0	制御領域		

【書類名】 図面

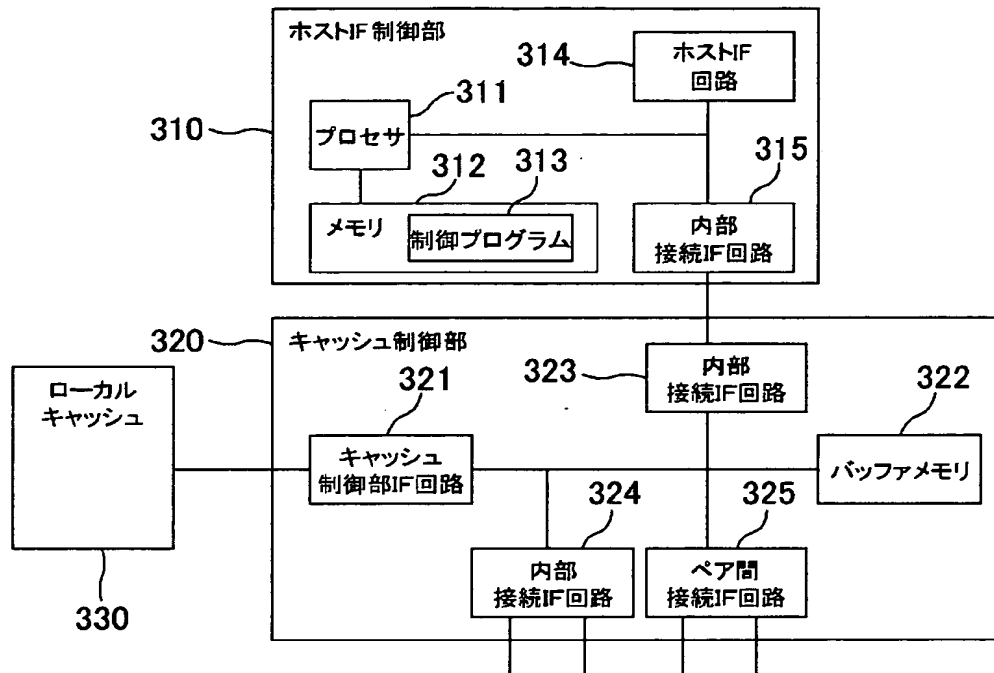
【図 1】



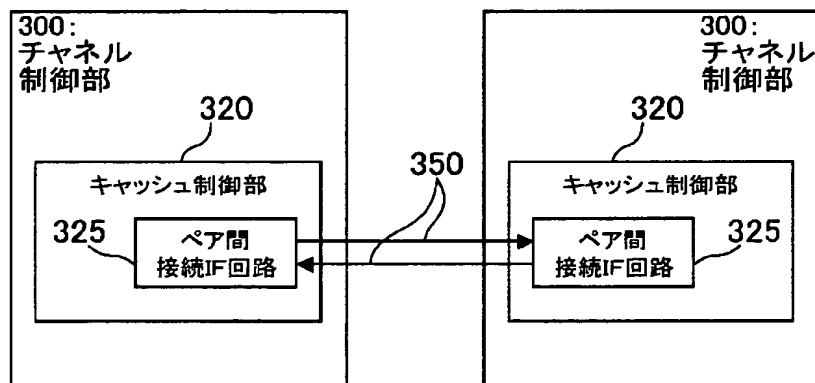
【図 2】



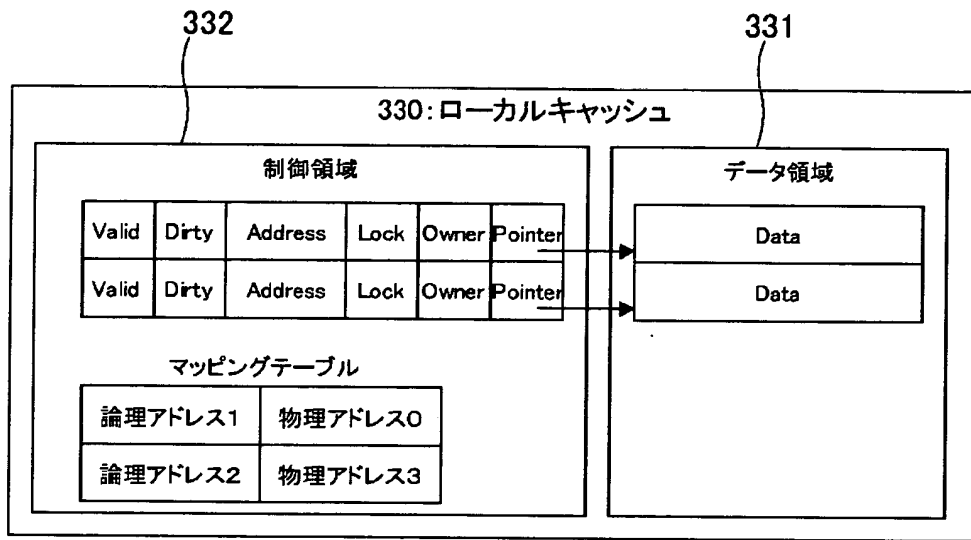
【図 3】



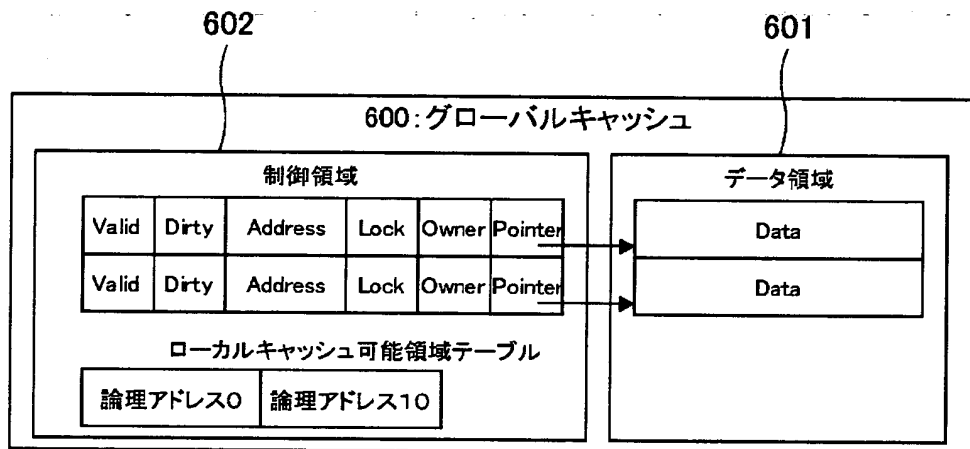
【図 4】



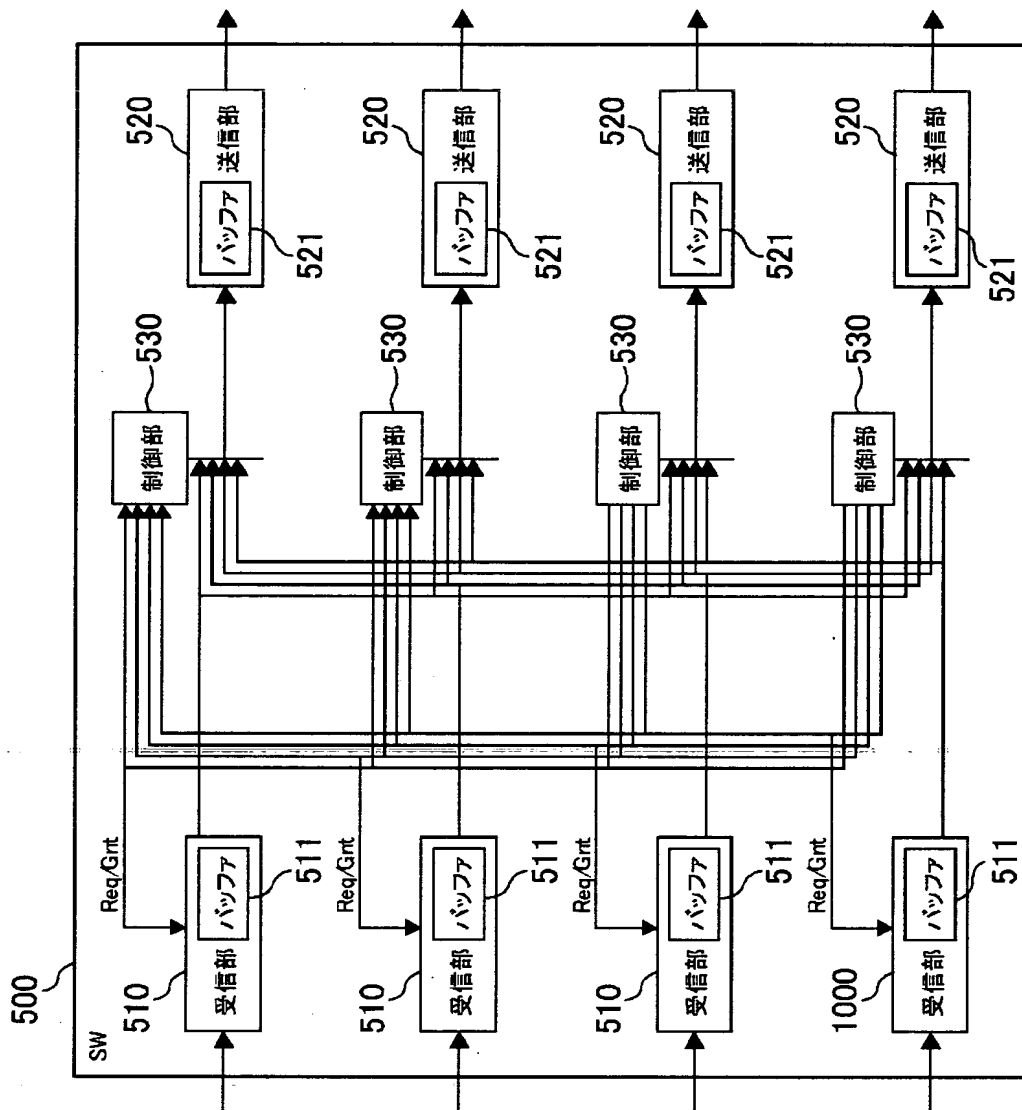
【図 5】



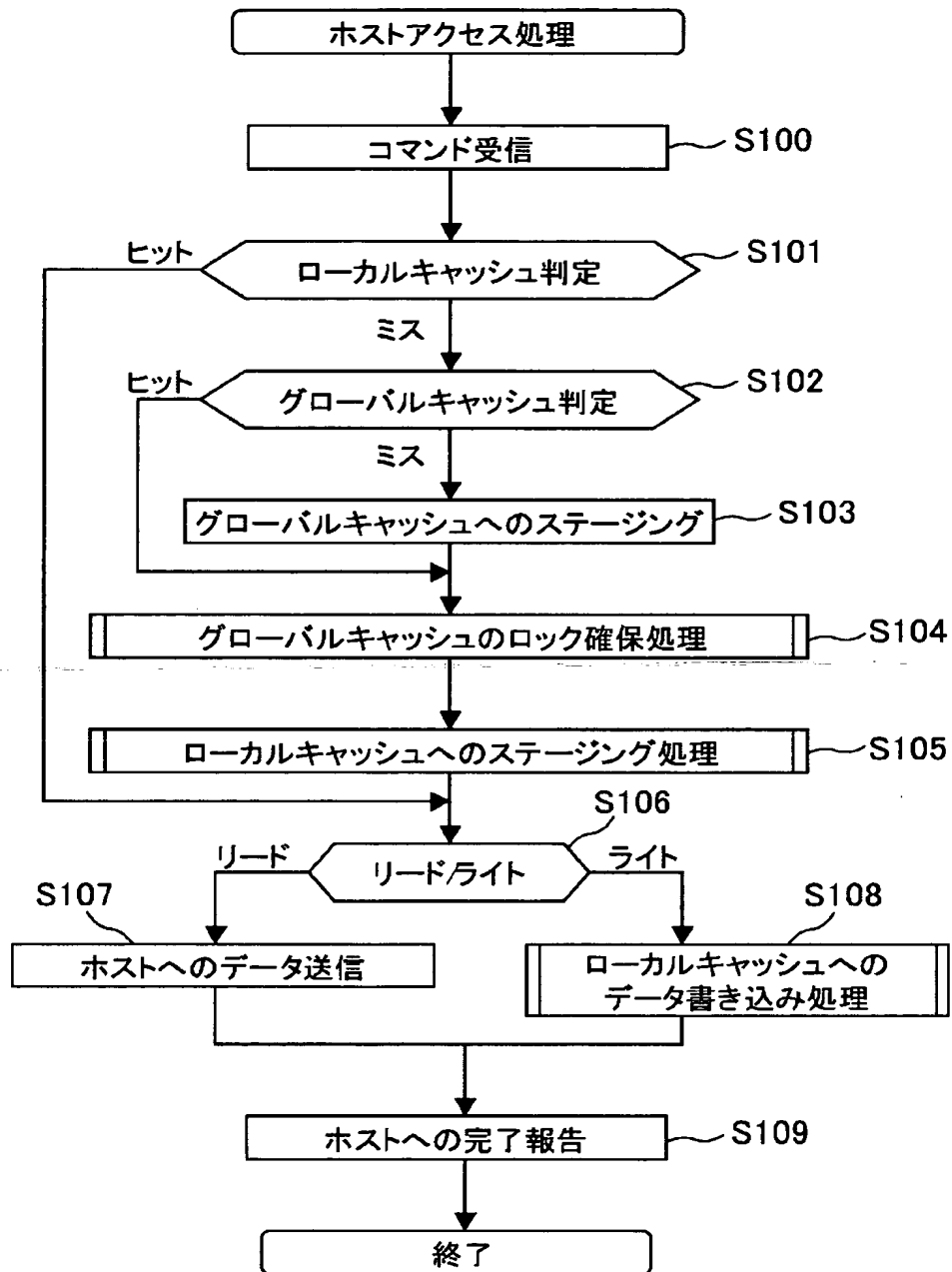
【図 6】



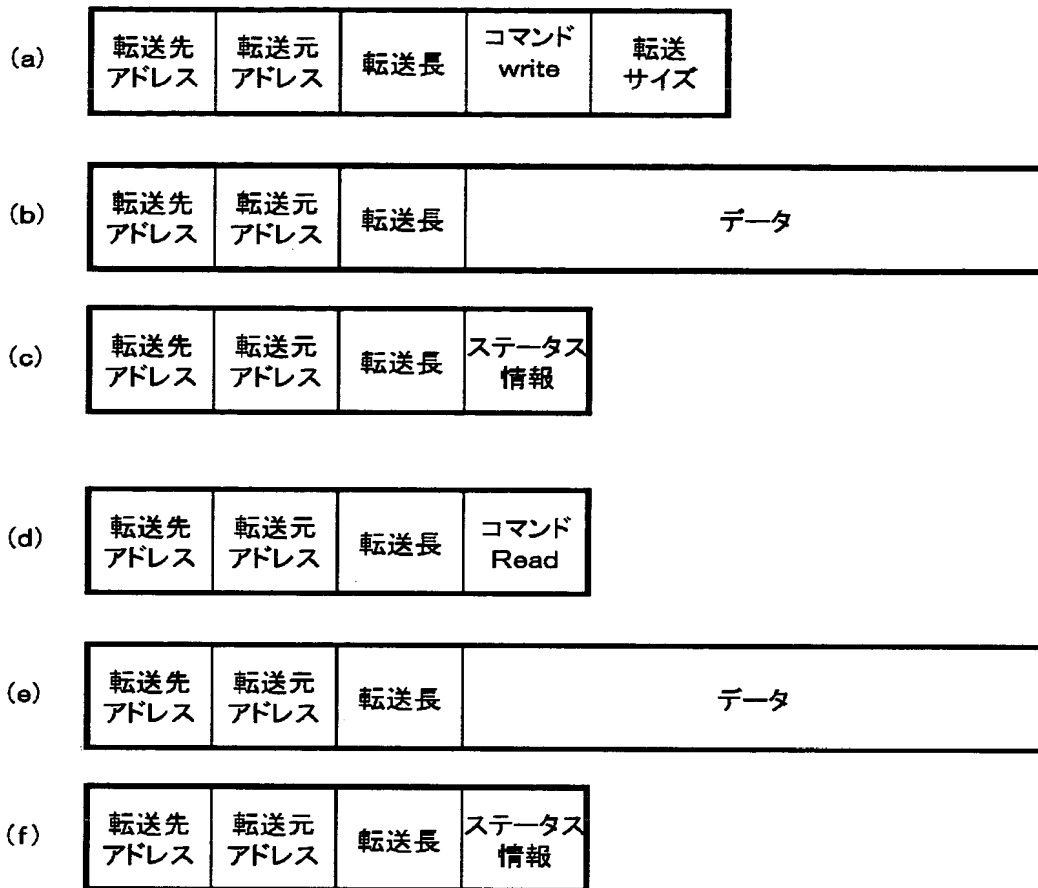
【图 7】



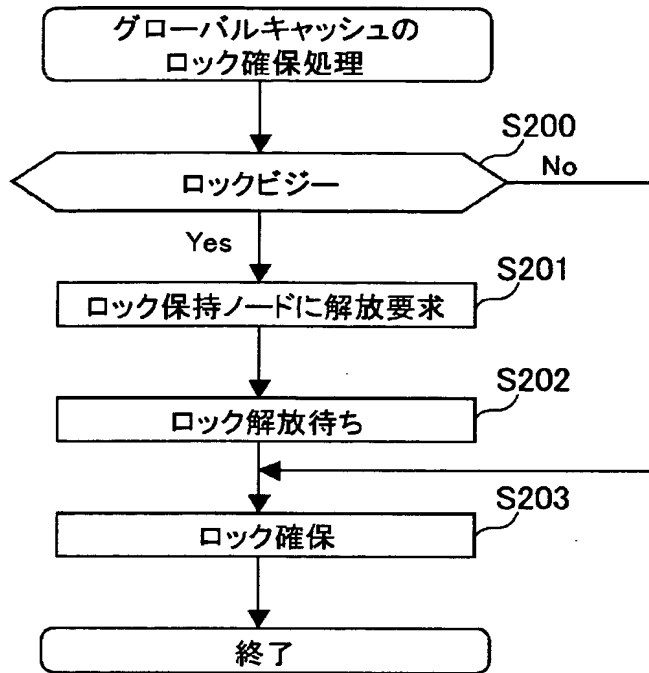
【図 8】



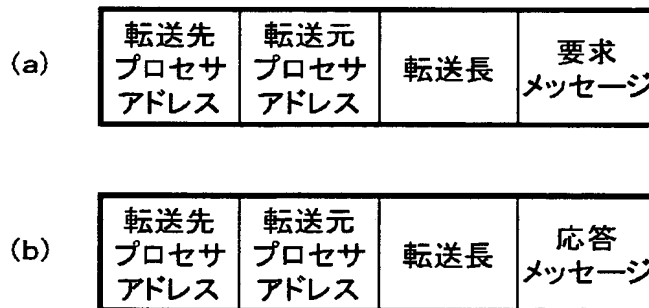
【図 9】



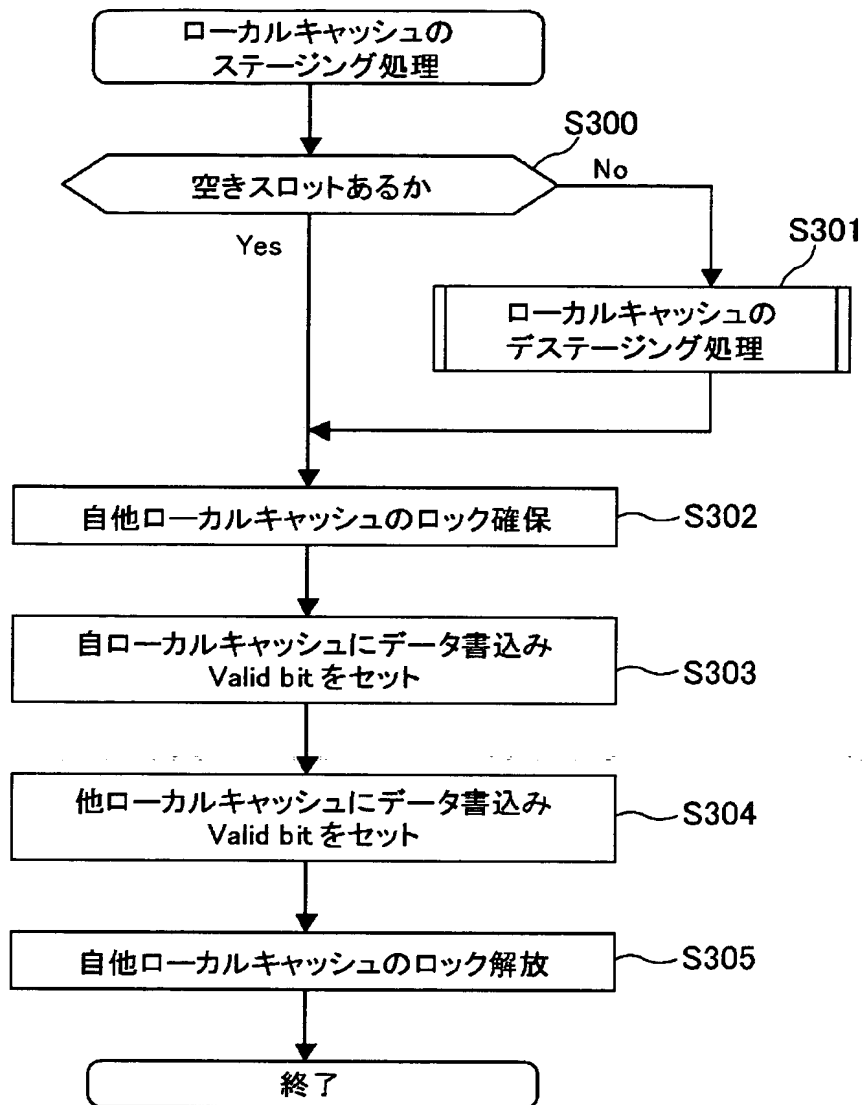
【図 1 0】



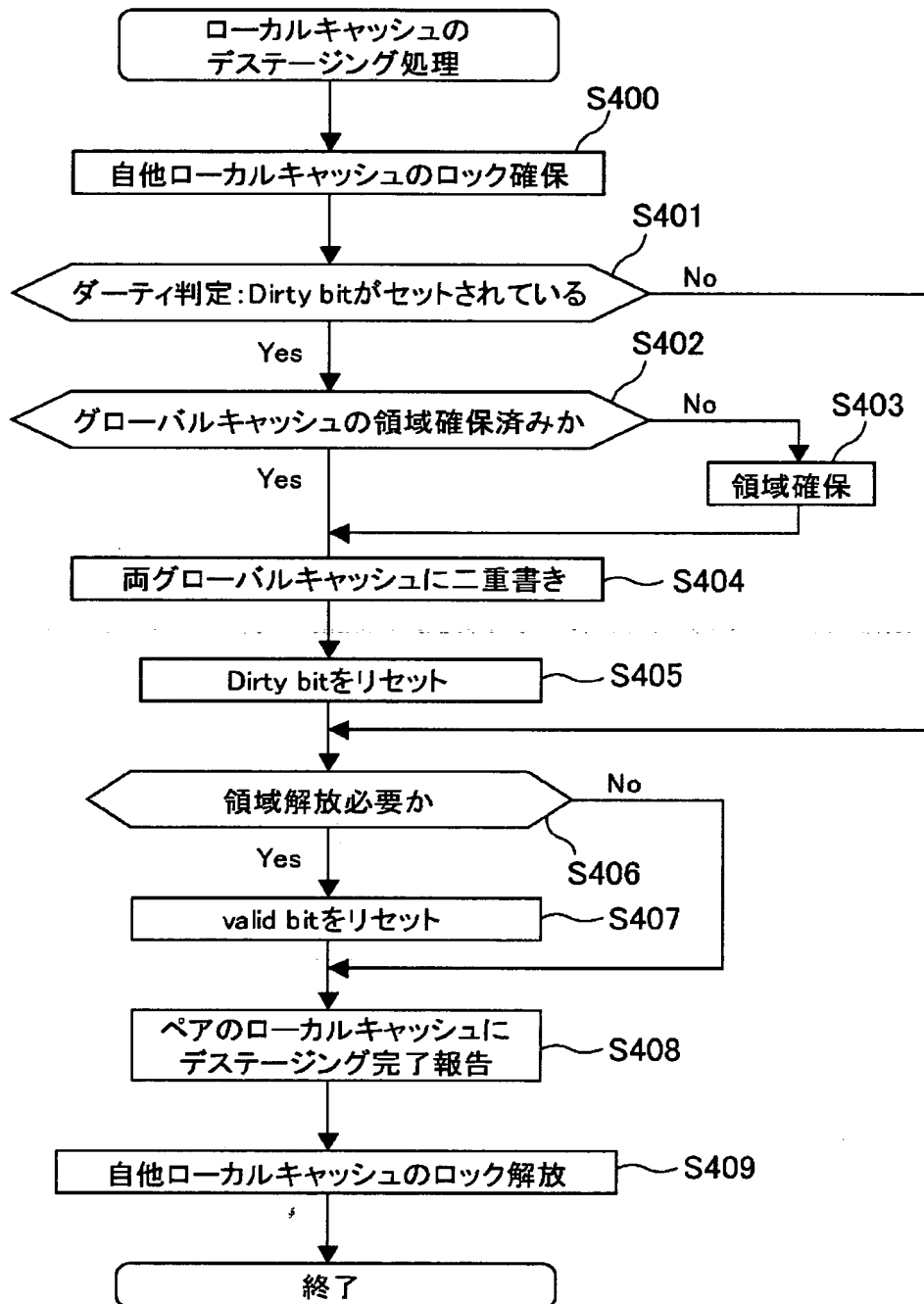
【図 1 1】



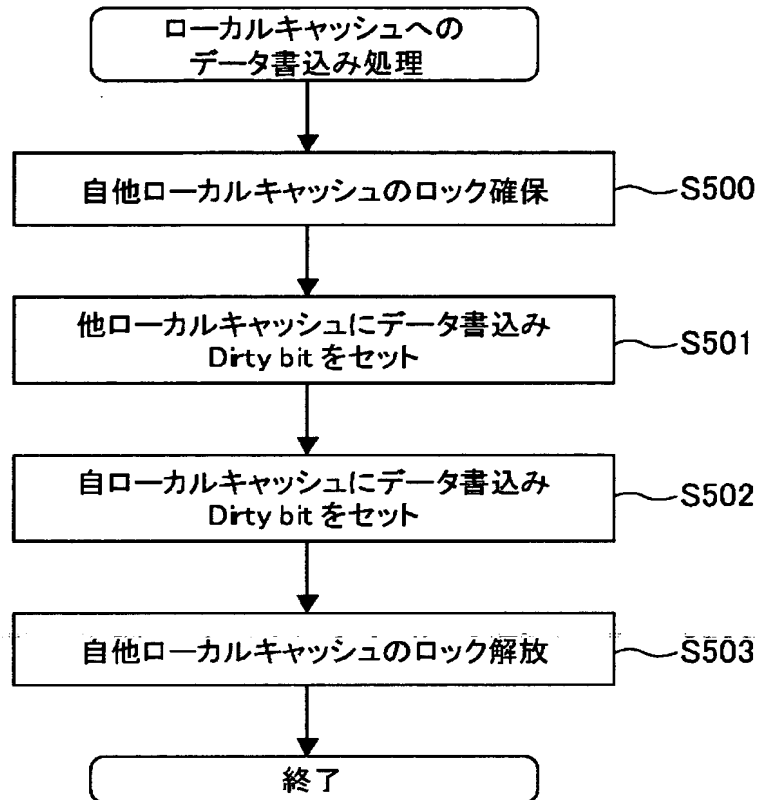
【図 1 2】



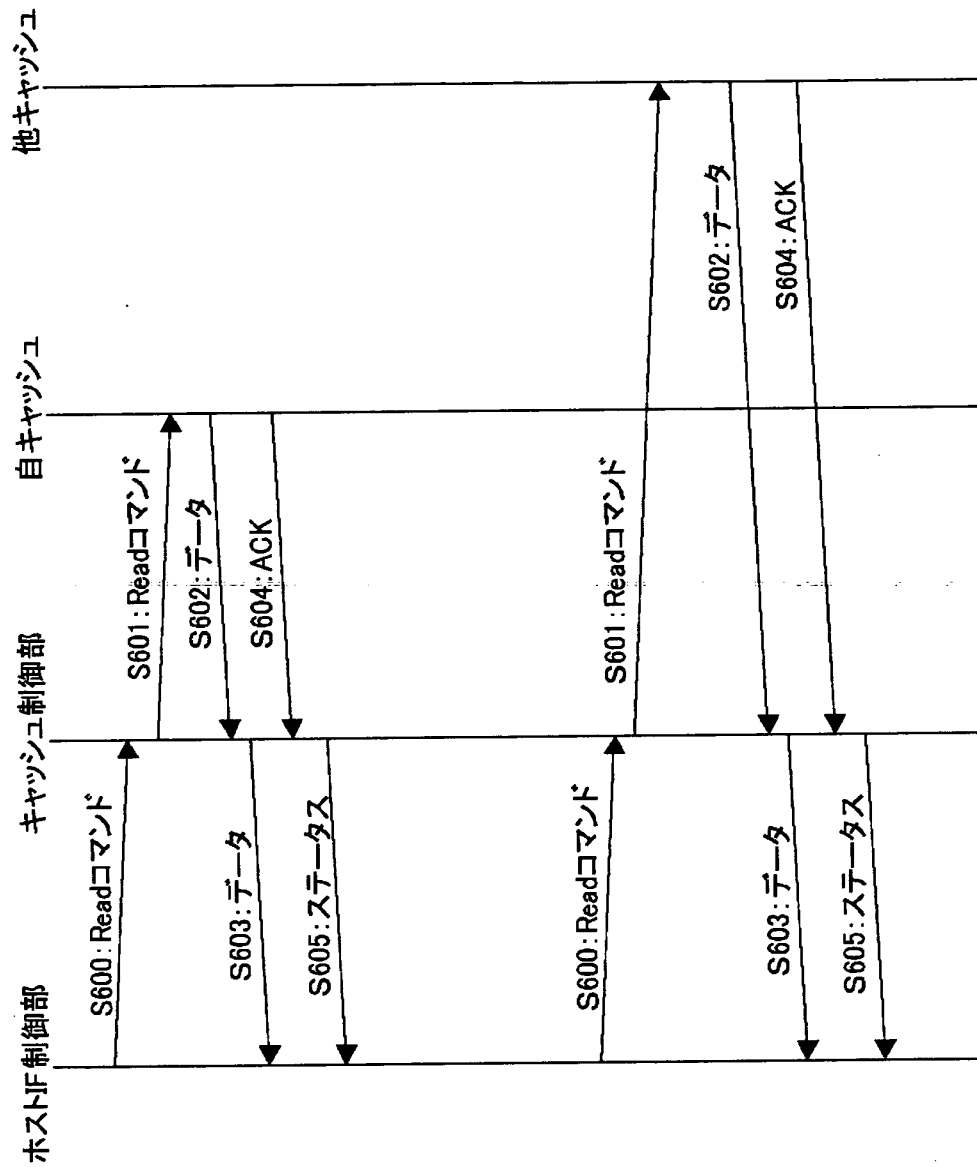
【図 1 3】



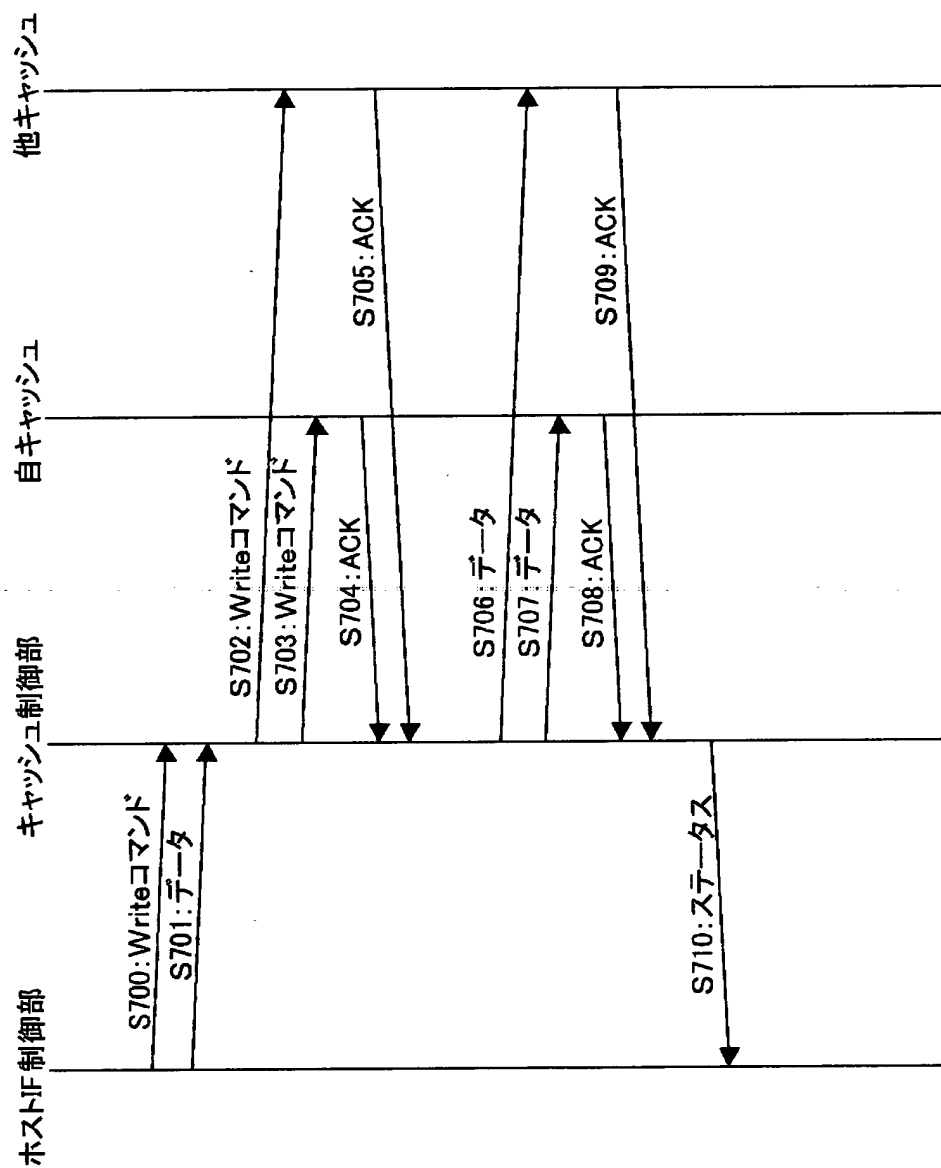
【図 1 4】



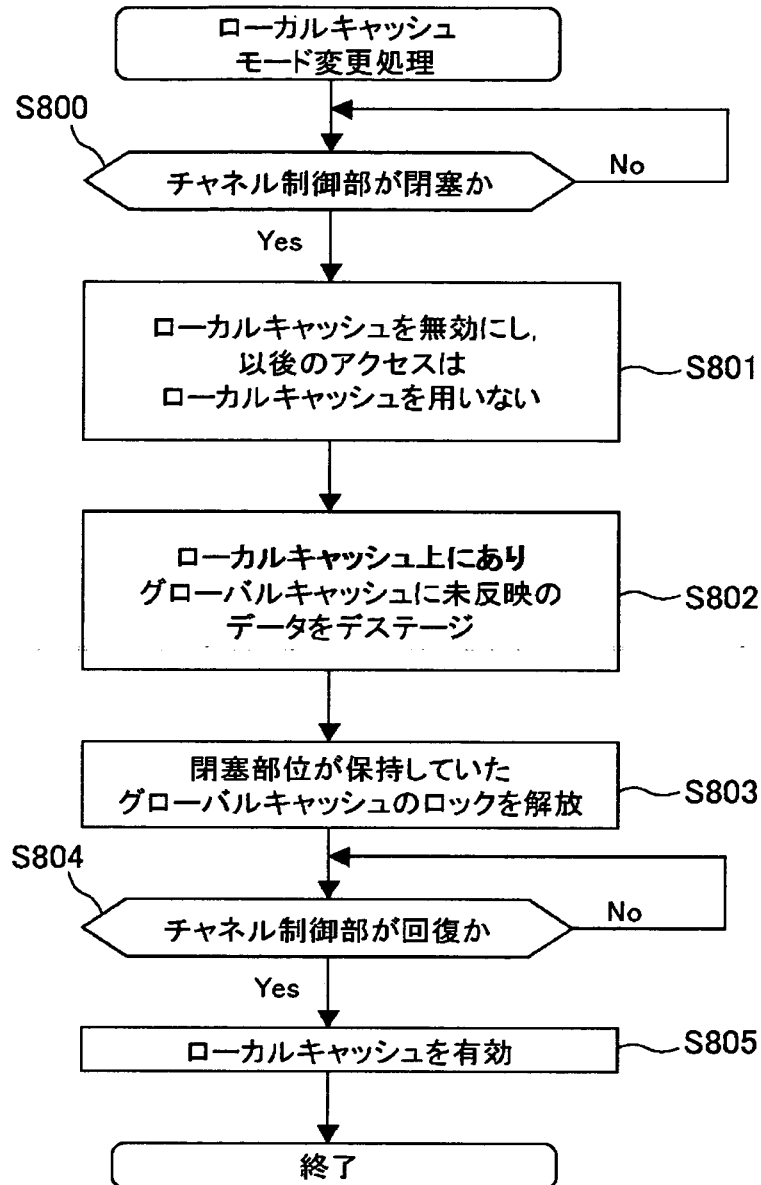
【図15】



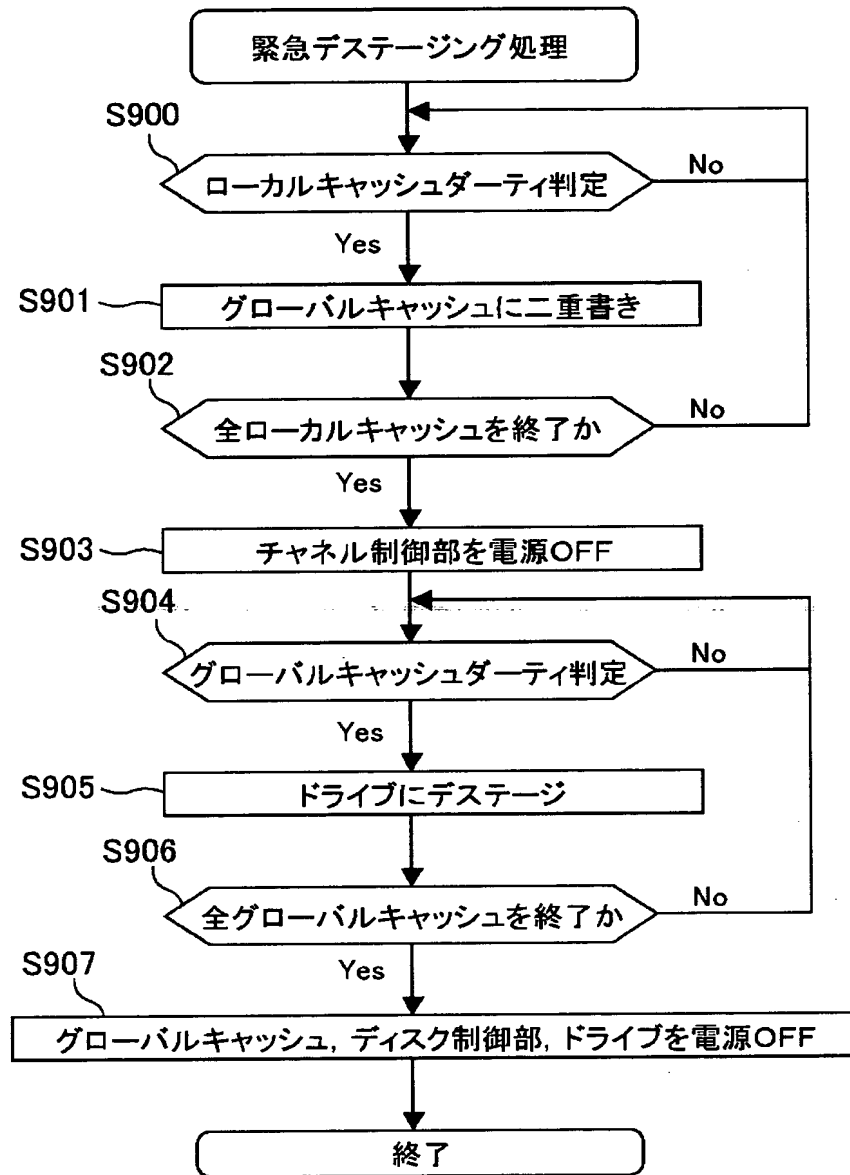
【図 16】



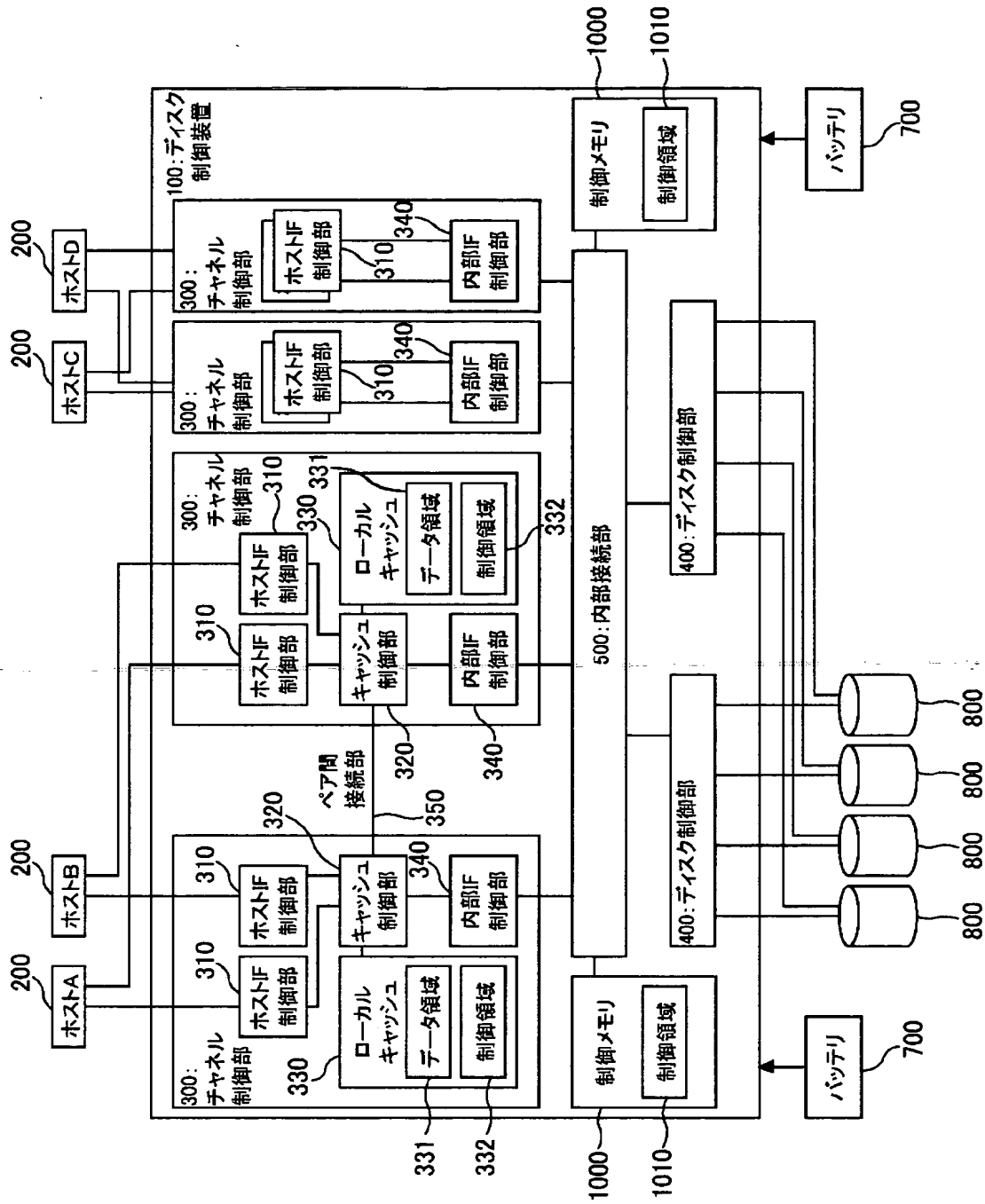
【図 1 7】



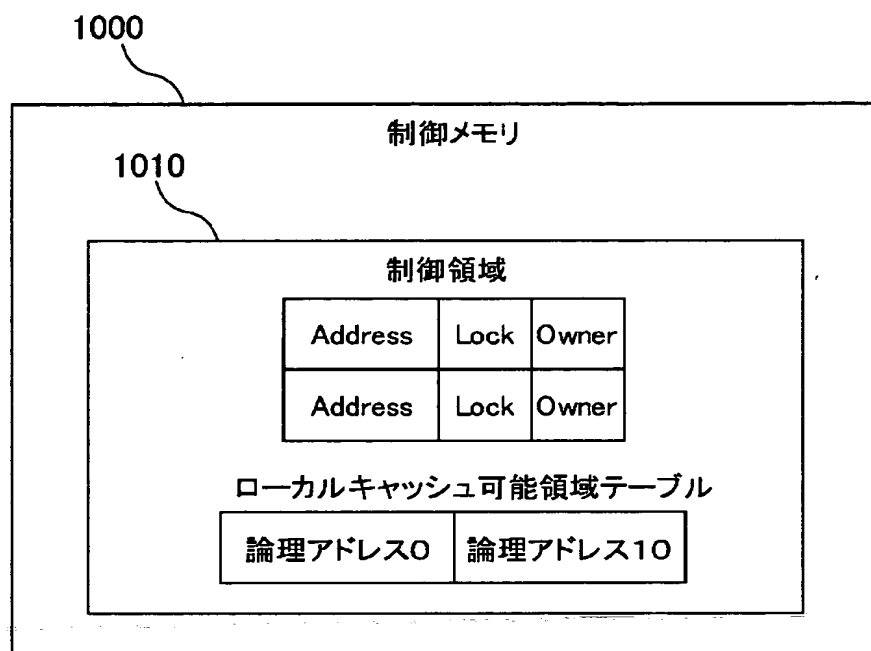
【図 1 8】



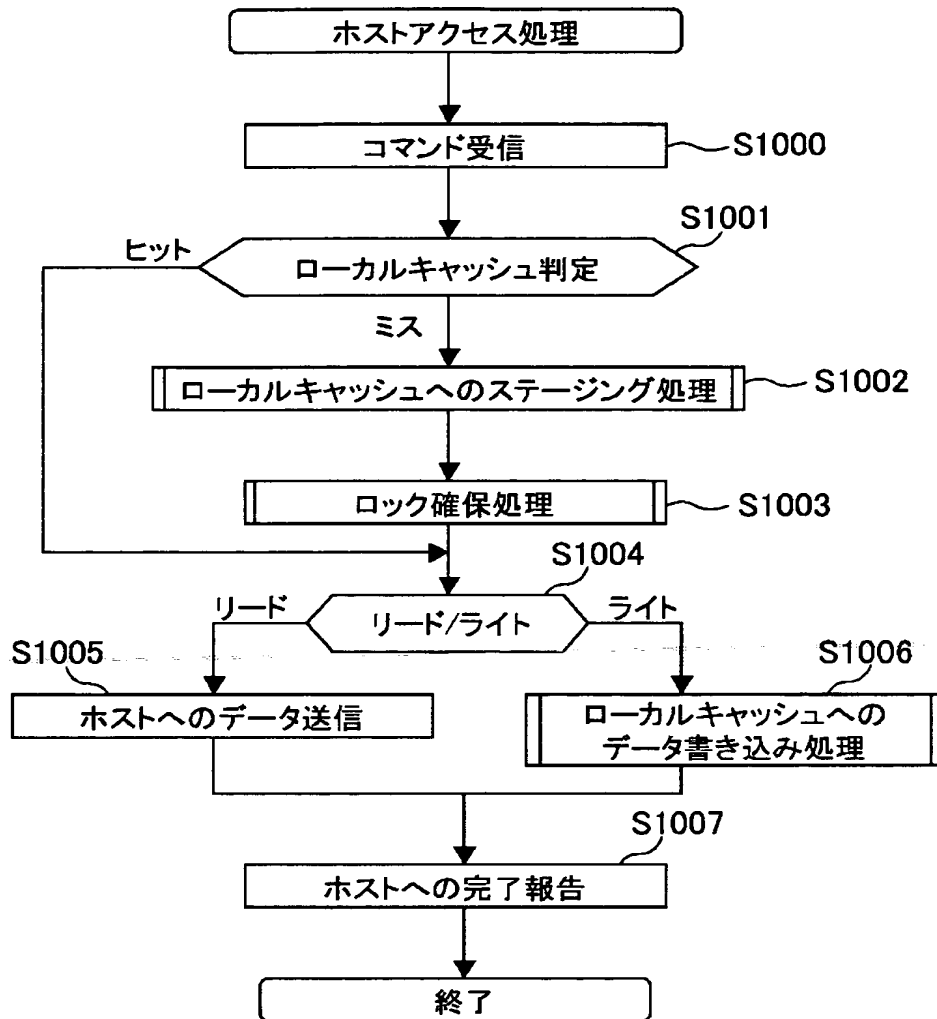
【図19】



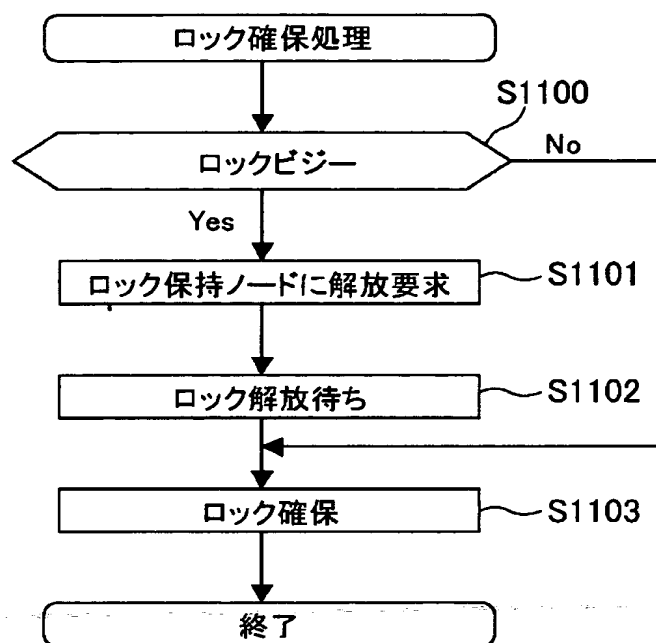
【図 2 0】



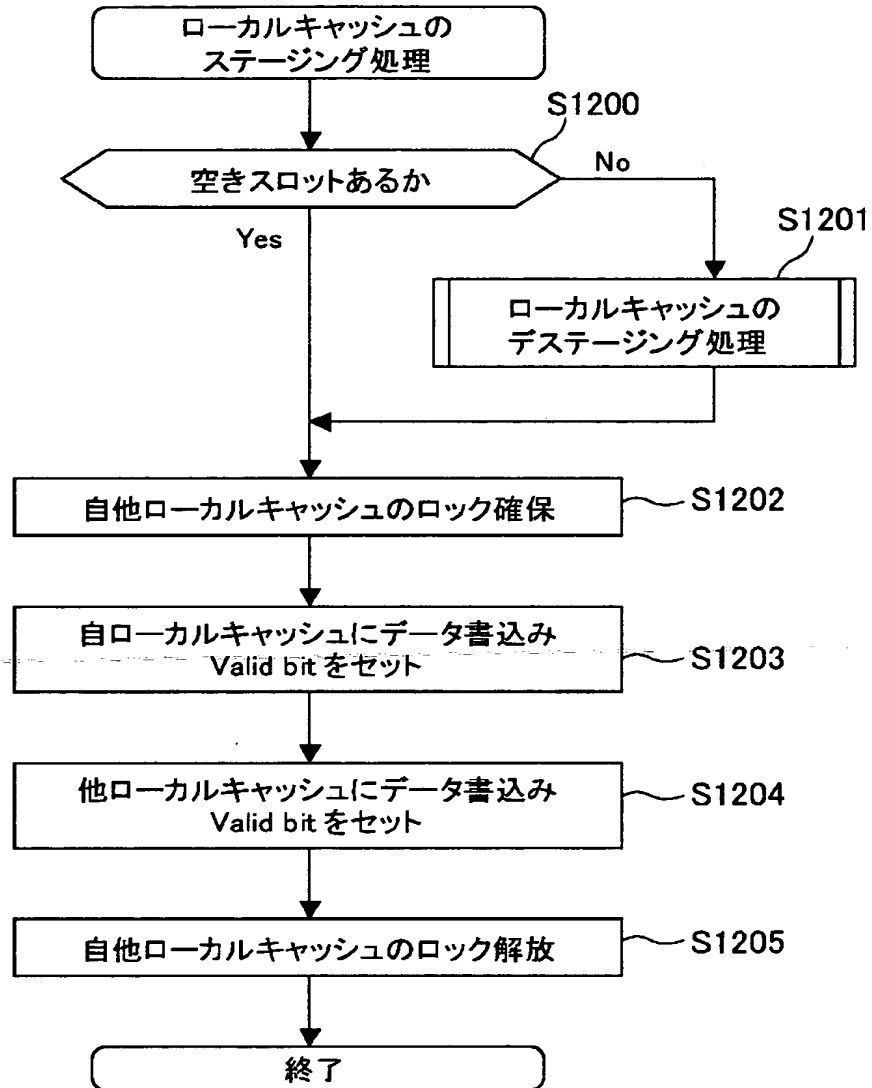
【図 21】



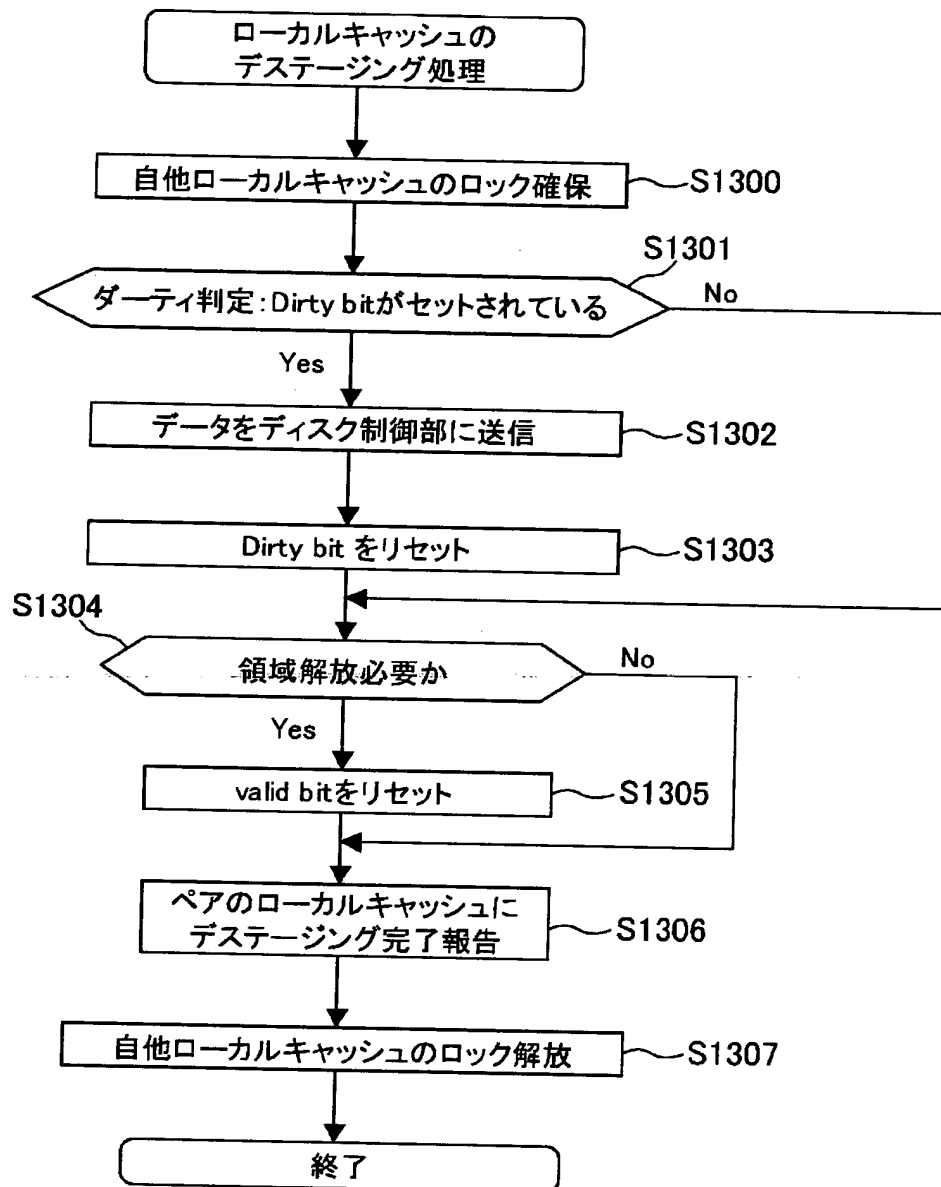
【図 2 2】



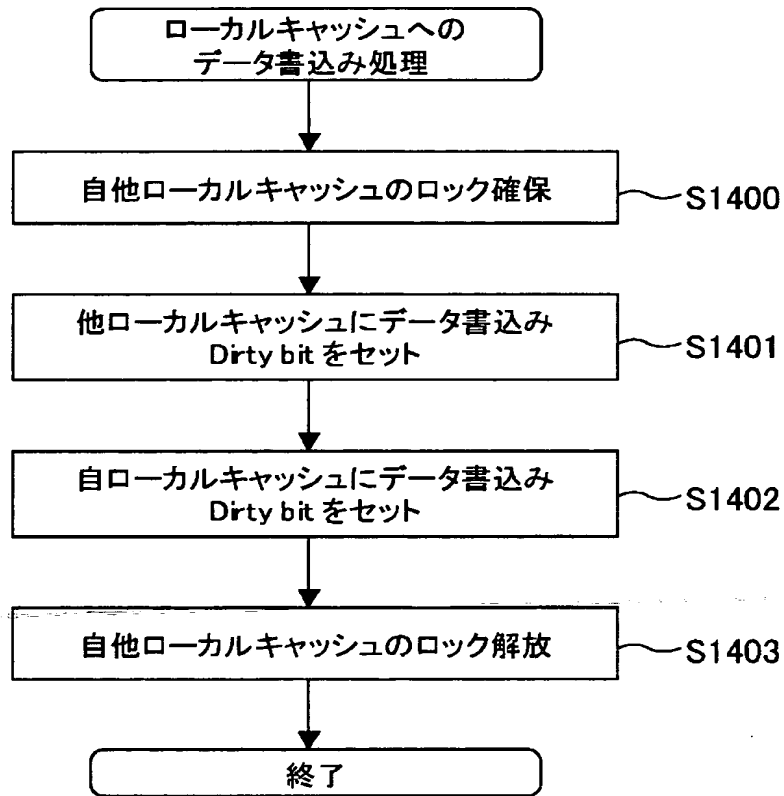
【図 2 3】



【図 2 4】



【図 2 5】



【書類名】 要約書

【要約】

【解決手段】 情報処理装置とのインタフェースを有する複数のチャンネル制御部と、データを記憶する記憶装置とのインタフェースを有するディスク制御部と、情報処理装置と記憶装置との間で授受されるデータを一時的に記憶するためのキャッシュメモリと、チャンネル制御部とディスク制御部とを相互に接続する内部接続部とを備える記憶制御装置であって、キャッシュメモリが、キャッシュメモリに記憶されたデータを相互に記憶するための専用のデータ転送路で相互に結合されているチャンネル制御部に配置されている。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日 1 9 9 0 年 8 月 3 1 日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所